

Algebra lineare *for dummies*

Sergio Polini

18 ottobre 2011

Indice

Premessa	2
1 Spazi vettoriali	3
1.1 Definizione	3
1.2 Sottospazi vettoriali	3
1.3 Indipendenza lineare e basi	6
1.4 Basi e coordinate	8
1.5 Somma e intersezione di sottospazi	10
2 Applicazioni lineari e matrici associate	11
2.1 Applicazioni lineari	11
2.2 Nucleo e immagine di un'applicazione lineare (1)	13
2.3 Matrici associate ad applicazioni lineari	14
2.3.1 Esempi di matrici associate (1)	16
2.3.2 Esempi di matrici associate (2)	18
2.4 Nucleo e immagine di un'applicazione lineare (2)	19
2.4.1 Trovare il nucleo	20
2.4.2 Trovare l'immagine	21
2.5 Teorema della nullità e del rango	22
2.6 Matrici di cambiamento di base	23
2.7 Matrici simili (primi cenni)	26
3 Autovalori, autovettori, autospazi	28
3.1 Le definizioni e il loro senso	28
3.2 Autovalori e polinomi caratteristici	30
3.3 Autovettori e autospazi	33
3.4 Diagonalizzazione di un operatore lineare	35

Premessa

Qualche tempo fa ho sostenuto un esame di algebra lineare e, mentre mi preparavo, seguivo il forum di www.matematicamente.it.

Mi stupiva vedere come tanti che ponevano domande avessero idee maledettamente confuse. Forse in alcuni corsi di laurea si chiede solo una ‘infarinatura’ di algebra lineare e le inevitabili semplificazioni rendono in realtà ancora più difficile la comprensione dei vari argomenti. Forse è anche difficile chiarire un dubbio sfogliando un testo diverso dal proprio, perché i testi di algebra lineare seguono ciascuno una propria impostazione, anche con differenze molto marcate (ci sono perfino titoli chiaramente contrapposti come *Linear algebra done right* e *Linear algebra done wrong*).

Probabilmente incide anche la stramaledetta faccenda dei semestri e degli esoneri, che a me proprio non piacciono. I semestri si riducono a tre mesi di lezione, che sono veramente troppo pochi. Non solo i docenti sono costretti a procedere speditamente, ma gli studenti non hanno il tempo di ‘digerire’ i vari argomenti, di capire il senso di una definizione o di una dimostrazione, di cogliere i nessi tra argomenti diversi. Gli esoneri poi costringono a dedicarsi alla manualità dei calcoli, che è certo necessaria ma non sufficiente, trascurando la comprensione della teoria.

Mi è sembrato di trovarne conferma nell’andamento del mio esame orale. Sono stato l’ultimo e ho quindi potuto assistere agli esami di tanti ragazzi. Ho visto molti di loro, che pure avevano superato lo scritto, naufragare su domande semplicissime, come la richiesta di una definizione (su ‘indipendenza lineare’ e ‘operatore diagonalizzabile’ gli scivoloni più frequenti).

E così mi è venuta l’idea di scrivere sul forum una serie di messaggi in un topic intitolato *Algebra lineare for dummies*: niente riduzioni a gradini, soluzione di sistemi di equazioni o calcolo di determinanti (queste cose in genere si sanno fare), quasi niente dimostrazioni formali (bene o male quelle nel libro di testo ci sono), ma solo il tentativo di illustrare il ‘senso’ dei primi argomenti, quelli che è fondamentale assimilare bene.

Con mia sorpresa, quella serie di messaggi ha avuto un discreto successo. Ancora oggi, nonostante siano passati ormai più di due anni, mi capita di sentirmi ringraziare, e proprio per l’obiettivo che volevo raggiungere: quelle quattro chiacchiere (che alcuni chiamano addirittura ‘dispense’!) «mi hanno permesso di capire molti concetti poco chiari e superare l’esame con il massimo dei voti», «sul mio libro certe cose non sono chiare ma grazie a queste diventano limpide».

Non si può certo restare indifferenti... Ho quindi deciso di ascoltare quelli che mi chiedevano di trasformare il tutto in un unico file e ho riportato quanto avevo scritto in quel topic.

Per parte mia, non posso non ringraziare i professori Alessandro Silva e Paolo Papi, che mi hanno insegnato quello che so. Due personalità molto diverse, ma due ottimi insegnanti. E anche il mitico (per chi ha avuto la fortuna di partecipare alle sue esercitazioni) Mario Marietti, al quale ho rubato un paio di esempi.

Ringrazio anche l’utente franced del forum. Le sue risposte sono sempre sintetiche, quasi stitiche, ma non perde un colpo e la sua sicurezza nel risolvere esercizi e problemi e quei suoi commenti stringatissimi mi hanno insegnato molto (dopo averci spesso dovuto riflettere per un po’...).

Ovviamente loro sono matematici, io un pivello. Ringrazio quindi (in ordine cronologico) gli utenti del forum che mi hanno corretto: alvinlee88 e Zcheggia mi hanno fatto notare che si deve parlare di somma e intersezione di sottospazi, non genericamente di spazi vettoriali; Fioravante Patrone ha sottolineato l'importanza di precisare che una base non è semplicemente un insieme di vettori, ma un insieme ordinato di vettori; dannoman1988 mi ha segnalato una frase che, scritta com'era, non aveva molto senso mentre salemgold ha trovato un errore in un esempio; simo90, Zilvius, raffamaiden e harm_burst mi hanno aiutato correggere varie sviste minori. È finita? Non credo. Cercherò quindi di tenere aggiornato il file quando arriveranno altre correzioni.

Grazie a tutti.

1 Spazi vettoriali

1.1 Definizione

Qualsiasi testo propone una definizione di *spazio vettoriale*, più o meno articolata. Vorrei proporne una semplice.

Spazio vettoriale. *Un insieme di elementi, detti vettori, che possono essere sommati tra loro e moltiplicati per uno scalare in modo tale che la somma di due vettori o la moltiplicazione di un vettore per uno scalare siano ancora elementi dell'insieme.*

Per *vettore* si intende un'entità composta: una n -upla di numeri, una matrice $m \times n$ (quindi $m \times n$ numeri), un polinomio (una somma di monomi), ma anche funzioni o altro. Anche singoli numeri possono essere vettori, ma si tratta solo di un caso particolare. I vettori si chiamano così per motivi storici, ma viene detto vettore anche l'elemento di uno spazio di polinomi o di matrici.

Per *scalare*, al contrario, si intende un singolo numero, generalmente reale o complesso. Nell'algebra lineare per dummies gli scalari sono solo numeri reali.

Dire che i vettori possono essere sommati vuol dire che valgono le normali proprietà della somma:

- proprietà associativa: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$;
- proprietà commutativa: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
- esistenza dello zero, detto *vettore nullo*: $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$;
- esistenza degli opposti: $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

Da notare che lo zero è necessario perché si possa sottrarre oltre che sommare. La sottrazione è infatti definita come addizione dell'opposto: $\mathbf{u} - \mathbf{w} = \mathbf{u} + (-\mathbf{w})$, dove $-\mathbf{w}$ è quel numero tale che $\mathbf{w} + (-\mathbf{w}) = \mathbf{0}$. Quindi *non può esserci uno spazio vettoriale senza un vettore nullo*.

1.2 Sottospazi vettoriali

Sottospazio vettoriale. *Un sottoinsieme di uno spazio vettoriale che conserva le stesse operazioni.*

Uno spazio vettoriale, ripetiamolo, è un insieme di elementi che possono essere sommati tra loro e moltiplicato per uno scalare *in modo tale che la somma di due*

vettori o la moltiplicazione di un vettore per uno scalare siano ancora elementi dell'insieme.

Un sottospazio vettoriale non è altro che un sottoinsieme di uno spazio che rispetti la stessa condizione; ad esempio, un sottoinsieme W di uno spazio vettoriale V tale che, presi due suoi elementi, la loro somma non appartiene più a W non è uno spazio vettoriale.

Problema. Determinare se un dato sottoinsieme W di uno spazio vettoriale V è un sottospazio, cioè se è a sua volta uno spazio vettoriale.

Soluzione. Per prima cosa si verifica che W contenga il vettore nullo (se non lo contiene abbiamo finito: non è un sottospazio), poi si cerca di capire se W comprende tutte le somme di suoi elementi e tutti prodotti di suoi elementi per uno scalare.

Esempio 1.1. L'insieme di tutte le terne di numeri reali è uno spazio vettoriale, che si indica con \mathbb{R}^3 . Infatti:

- a) contiene anche la terna $(0, 0, 0)$, il vettore nullo;
- b) la somma di due terne di numeri reali è ancora una terna di numeri reali, ad esempio:

$$(1, 2, 3) + (e, \pi, \sqrt{2}) = (1 + e, 2 + \pi, 3 + \sqrt{2})$$

- c) la moltiplicazione di una terna di numeri reali per un qualsiasi scalare (anch'esso un numero reale) è ancora una terna di numeri reali:

$$0.5 \cdot (1, 2, 3) = (0.5, 1, 1.5)$$

Esempio 1.2. L'insieme W di tutte le terne di numeri reali che abbiano come terzo componente il numero 1 è un sottoinsieme di \mathbb{R}^3 , ma non è un sottospazio. Infatti non contiene il vettore nullo, $(0, 0, 0) \notin W$, perché il terzo componente è diverso da 1. Fine.

Esempio 1.3. L'insieme delle matrici 2×2 è uno spazio vettoriale. Infatti:

- a) contiene il vettore nullo: $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$;

- b) la somma di matrici 2×2 è ancora una matrice 2×2 :

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a + e & b + f \\ c + g & d + h \end{bmatrix}$$

- c) il prodotto di una matrice 2×2 per uno scalare è ancora una matrice 2×2 :

$$2 \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2a & 2b \\ 2c & 2d \end{bmatrix}$$

Esempio 1.4. L'insieme delle matrici 2×2 con la seconda colonna pari al doppio della prima è un sottoinsieme del precedente ed è a sua volta uno spazio vettoriale. Infatti:

- a) contiene il vettore nullo ($2 \cdot 0 = 0$);

b) la somma di due matrici con la seconda colonna pari al doppio della prima è ancora una matrice dello stesso tipo:

$$\begin{bmatrix} a & 2a \\ b & 2b \end{bmatrix} + \begin{bmatrix} c & 2c \\ d & 2d \end{bmatrix} = \begin{bmatrix} a+c & 2a+2c \\ b+d & 2b+2d \end{bmatrix} = \begin{bmatrix} a+c & 2(a+c) \\ b+d & 2(b+d) \end{bmatrix}$$

c) il prodotto di una matrice con la seconda colonna pari al doppio della prima per uno scalare è ancora una matrice dello stesso tipo:

$$k \begin{bmatrix} a & 2a \\ b & 2b \end{bmatrix} = \begin{bmatrix} ka & k2a \\ kb & k2b \end{bmatrix} = \begin{bmatrix} ka & 2(ka) \\ kb & 2(kb) \end{bmatrix}$$

Esempio 1.5. L'insieme

$$W = \left\{ \begin{bmatrix} a+b \\ a-b \\ \frac{a^2+1}{a} \\ b \end{bmatrix} : a, b \in \mathbb{R} \right\}$$

è un sottoinsieme di \mathbb{R}^4 , ma non è uno spazio vettoriale. Si vede subito che contiene il vettore nullo (basta prendere $a = b = 0$), ma compaiono due operazioni strane (un quadrato e una divisione) che non rientrano nella definizione di spazio vettoriale. Invece di procedere come negli esempi precedenti, possiamo provare a cercare una scorciatoia, cioè un controesempio. Il primo componente di W è la somma del terzo e del quarto, e le somme non spaventano. Nel secondo abbiamo una divisione sospetta per $a^2 + 1$; proviamo quindi a creare un vettore \mathbf{v} sostituendo a con 1 e lasciando $b = 0$ (la cosa più semplice), e poi a vedere che succede se sommiamo \mathbf{v} a se stesso (cioè a moltiplicarlo per 2):

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1/2 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{w} = \mathbf{v} + \mathbf{v} = 2\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 0 \end{bmatrix}$$

Si vede che \mathbf{w} non appartiene allo spazio perché, se il terzo componente è 2 e il quarto è 0, il primo deve essere $2 + 0 = 2$ (e ci siamo), ma il secondo dovrebbe essere: $(2 - 0)/(2^2 + 1) = 2/5$ e non è così. Questo vuol dire che \mathbf{w} non appartiene a W , cioè che W non è uno spazio vettoriale: dovrebbe contenere la somma di due suoi vettori qualsiasi (anche uguali), nonché il prodotto di qualsiasi suo vettore per un qualsiasi scalare, ma abbiamo trovato un caso in cui ciò non succede.

Osservazione. I primi quattro esempi sono facili, il quinto un po' meno, ma illustra un aspetto importante. Se si pensa di poter dimostrare che un sottoinsieme W di uno spazio vettoriale V è un sottospazio, allora c'è un solo modo: sviluppare, con passaggi algebrici spesso semplici, i casi generali di somma e di moltiplicazione per uno scalare; si devono cioè usare simboli che possano valere per qualsiasi numero. Quando invece si cerca di dimostrare il contrario, è *sufficiente mostrare un solo caso*, usando normali numeri e passaggi aritmetici. Infatti, se una proposizione non vale anche per un solo caso, allora non può valere per la generalità dei casi. In altri termini:

- se voglio dimostrare che $(a+b)^2 = a^2 + 2ab + b^2$ non posso fare altro che svolgere i calcoli con le lettere, stando solo attento a non incappare in situazioni come la divisione per zero o l'estrazione di radice quadrata di un numero negativo;
- se voglio dimostrare che $(a+b)^2 \neq a^2 + b^2$, mi basta $(1+2)^2 \neq 1+2^2$ e ho finito.

1.3 Indipendenza lineare e basi

Uno spazio vettoriale comprende tutte le somme di suoi elementi e tutti i prodotti di suoi elementi per uno scalare. Quindi, se $\mathbf{u}, \mathbf{v}, \mathbf{w}$ sono suoi elementi, è un suo elemento anche una combinazione di somme e prodotti per scalari, come $3\mathbf{u} + 2\mathbf{v} - 1/2\mathbf{w}$. In generale, si ha la seguente definizione:

Combinazione lineare. *Dati n vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, e dati n scalari k_1, \dots, k_n , si dice combinazione lineare degli n vettori il vettore:*

$$\sum_{i=1}^n k_i \mathbf{v}_i = k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \dots + k_n \mathbf{v}_n$$

Osservazione. La definizione di spazio vettoriale può essere riformulata in termini di combinazioni lineari; un insieme V di vettori è uno spazio vettoriale se e solo se qualsiasi combinazione lineare di un sottoinsieme di vettori appartenenti a V è ancora un vettore appartenente a V .

Ne segue:

Vettori generatori. *Dati n vettori, l'insieme V di tutti i vettori che siano loro combinazioni lineari è uno spazio vettoriale (comprende, tra l'altro, il vettore nullo: basta che i k_i coefficienti siano tutti nulli). Un tale spazio viene detto generato da quegli n vettori, che a loro volta vengono detti generatori.*

Esempio 1.6. Consideriamo il normale piano cartesiano. Tutti i suoi punti hanno coordinate del tipo (x, y) , dove x è l'ascissa e y è l'ordinata. Lo spazio vettoriale \mathbb{R}^2 è analogo: dati i suoi vettori $(1, 0)$ e $(0, 1)$, qualsiasi suo elemento è del tipo (x, y) , e questa altro non è che una combinazione lineare di quei due vettori: $(x, y) = x(1, 0) + y(0, 1)$. In questo senso i due vettori generano tutto \mathbb{R}^2 : non esiste un elemento di \mathbb{R}^2 che non possa essere espresso come combinazione lineare di quei due.

Si pone ora un problema: dato uno spazio vettoriale, quanti generatori sono necessari?

«Anche infiniti» non sarebbe una risposta molto pratica. È evidente (anche se vedremo presto un motivo preciso) che è molto più interessante trovare il *numero minimo* di generatori necessari. Per farlo, è necessario introdurre la seguente definizione:

Indipendenza lineare. *Dato uno spazio vettoriale V e dati suoi n vettori, questi si dicono linearmente dipendenti se esistono n scalari non tutti nulli tali che:*¹

$$k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \dots + k_n \mathbf{v}_n = \mathbf{0}_V$$

Altrimenti si dicono linearmente indipendenti.

¹Si indica con $\mathbf{0}_V$ il vettore nullo dello spazio vettoriale V . 0 è un numero, $\mathbf{0}_V$ può essere qualsiasi cosa: una matrice, un polinomio ecc.

Che vuol dire? Immaginiamo che k_1 sia diverso da 0 e che si abbia:

$$k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \cdots + k_n \mathbf{v}_n = \mathbf{0}_V$$

Essendo $k_1 \neq 0$ posso dividere per k_1 ottenendo:

$$\mathbf{v}_1 = -\frac{k_2}{k_1} \mathbf{v}_2 - \frac{k_3}{k_1} \mathbf{v}_3 - \cdots - \frac{k_n}{k_1} \mathbf{v}_n$$

cioè potrei esprimere \mathbf{v}_1 come combinazione lineare degli altri. In questo senso \mathbf{v}_1 viene detto linearmente dipendente dagli altri: non aggiunge nulla, è ‘solo’ una combinazione lineare di altri vettori (quindi non serve come generatore: se gli altri generano \mathbf{v}_1 , possono generare tranquillamente qualsiasi altro vettore nella cui generazione \mathbf{v}_1 intervenga; lo possono sostituire). Se invece posso ottenere $k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \cdots + k_n \mathbf{v}_n = \mathbf{0}_V$ solo con scalari tutti nulli, non posso dividere per nessuno e quindi non posso esprimere nessuno dei vettori come combinazione lineare degli altri; in questo senso ciascun vettore è linearmente indipendente dagli altri (che non lo possono sostituire nella generazione di altri vettori; è un generatore necessario).

Vi sono alcuni casi particolari:

- a) un vettore singolo non nullo \mathbf{v} è sempre linearmente indipendente, in quanto $k\mathbf{v} = \mathbf{0}_V$ richiede ovviamente $k = 0$;
- b) il vettore nullo è invece linearmente dipendente, in quanto $k\mathbf{0}_V = \mathbf{0}_V$ per un qualsiasi k ;
- c) due vettori non nulli sono linearmente dipendenti se e solo se sono proporzionali; se infatti $\mathbf{w} = 2\mathbf{v}$, per avere $k_1 \mathbf{v} + k_2 \mathbf{w} = \mathbf{0}$ basta prendere $k_1 = -2$ e $k_2 = 1$: $-2\mathbf{v} + \mathbf{w} = -2\mathbf{v} + (2\mathbf{v}) = \mathbf{0}_V$;
- d) due o più vettori di cui almeno uno nullo costituiscono un insieme linearmente dipendente; basta infatti che non sia 0 il coefficiente di un vettore nullo:

$$k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \cdots + k_n \mathbf{0}_V = \mathbf{0}_V \quad \Rightarrow \quad \mathbf{0}_V = -\frac{k_1}{k_n} \mathbf{v}_1 - \frac{k_2}{k_n} \mathbf{v}_2 - \cdots - \frac{k_{n-1}}{k_n} \mathbf{v}_{n-1}$$

Come si fa a capire se alcuni vettori sono linearmente indipendenti? Ci sono vari modi, ma visto che in genere sono tutti in grado di ridurre a gradini o calcolare un determinante passo oltre. Dato un insieme di vettori, come si fa a trovare un suo sottoinsieme costituito solo da vettori linearmente indipendenti? Può essere affrontato con le stesse tecniche di sopra; passo oltre.

Quanto visto finora consente, finalmente, di chiarire cosa si intende per *base* di uno spazio vettoriale e per *coordinate* di un vettore rispetto a una data base.

Base. Dato uno spazio vettoriale V , **una** sua base è un insieme ordinato di suoi vettori linearmente indipendenti capaci di generare tutto V .

Il numero degli elementi di una base di V viene detto *dimensione* di V .

Attenzione! Si dice sempre **una base**, mai **la base**, perché le basi di uno spazio vettoriale sono infinite.

Esempio 1.7. Torniamo a \mathbb{R}^2 , spazio vettoriale rappresentabile come un normale piano cartesiano. I vettori $(1, 0)$ e $(0, 1)$ ne costituiscono una base, perché, come

visto, qualsiasi vettore (x, y) può essere espresso come loro combinazione lineare. Ad esempio:

$$(4, 8) = 4(1, 0) + 8(0, 1)$$

Ma non è certo l'unica base: se prendo i vettori $(2, 0)$ e $(0, 2)$, avrò:

$$(4, 8) = 2(2, 0) + 4(0, 1)$$

Cambiano i coefficienti, ma $(4, 8)$ è combinazione lineare anche degli elementi della nuova base. E così via: possono essere basi $(12, 0)$ e $(0, -3)$, $(\pi, 0)$ e $(0, e)$ ecc. Non solo: se due vettori \mathbf{v}, \mathbf{w} costituiscono una base, cambiando il loro ordine si ottiene una base diversa (ricordiamolo: una base è un insieme *ordinato* di vettori).

Perché mai i vettori che sono elementi di una base devono essere linearmente indipendenti? Per capirlo, si deve introdurre un'altra definizione.

1.4 Basi e coordinate

Coordinate. *Dati uno spazio vettoriale V , una sua base $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ e un vettore \mathbf{v} , si dicono coordinate di \mathbf{v} i coefficienti che esprimono \mathbf{v} come combinazione lineare degli elementi di B .*

Esempio 1.8. Nello spazio vettoriale \mathbb{R}^2 , rispetto alla base $\{(1, 0), (0, 1)\}$ le coordinate di $(4, 8)$ non sono altro che $(4, 8)$. Rispetto alla base $\{(2, 0), (0, 2)\}$, le coordinate sono $(2, 4)$. Notare che, *mentre il vettore rimane lo stesso, le sue coordinate cambiano se cambia la base*, cioè le coordinate di un vettore sono sempre coordinate rispetto a una base fissata.

Torniamo alla domanda: perché mai i vettori che sono elementi di una base devono essere linearmente indipendenti?

Risposta: perché se così non fosse non sarebbe possibile trovare coordinate univoche per un vettore.

Esempio 1.9. Proseguiamo sulla linea dell'esempio precedente. Immaginiamo per assurdo che una base sia $B = \{(1, 0), (0, 1), (1, 1)\}$; è evidente che ora i tre vettori non sono più linearmente indipendenti (il terzo non è altro che la somma dei primi due). Proviamo a trovare le coordinate del vettore $(4, 8)$ rispetto a B :

$$\begin{aligned} (4, 8) &= 4(1, 0) + 8(0, 1) + 0(1, 1), & \text{coordinate: } (4, 8, 0) \\ (4, 8) &= 0(1, 0) + 4(0, 1) + 4(1, 1), & \text{coordinate: } (0, 4, 4) \\ (4, 8) &= 5(1, 0) + 9(0, 1) - 1(1, 1), & \text{coordinate: } (5, 9, -1) \\ (4, 8) &= 3(1, 0) + 7(0, 1) + 1(1, 1), & \text{coordinate: } (3, 7, 1) \\ &\text{ecc.} \end{aligned}$$

Come scegliere?

Morale: *non si può capire cosa è una base se non si hanno le idee chiare sull'indipendenza lineare; non si può capire a cosa serve una base se non si hanno le idee chiare sulle coordinate.*

Esempio 1.10. Cambiamo argomento. Una possibile base per lo spazio vettoriale delle matrici 2×2 è:

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

Rispetto a essa, le coordinate di $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ sono $(3, 2, 1, 4)$. Facile. Se cambiamo l'ordine degli elementi della base otteniamo una base diversa, e questo ha conseguenze immediate sulle coordinate; se scegliamo come base:

$$\left\{ \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

le coordinate di $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ sono $(1, 2, 3, 4)$.

Osservazione. Anche quando uno spazio vettoriale è 'strano' (cioè normalissimo, ma percepito come strano da chi pensa solo a spazi \mathbb{R}^n), le coordinate dei suoi elementi sono n -uple di numeri reali, *sono cioè vettori di \mathbb{R}^n* . C'è dietro qualche aspetto teorico affascinante, a volte trascurato dai manuali studiati da coloro che... potrebbero aver bisogno di queste note. Per farla semplice:

- a) a ogni vettore (n -upla di numeri, matrice, polinomio ecc.) di uno spazio di dimensione n corrisponde una e una sola n -upla di coordinate rispetto a una base fissata, cioè un solo vettore di \mathbb{R}^n , e viceversa (ogni punto del piano cartesiano ha una sola ascissa e una sola ordinata; a una data ascissa e a una data ordinata corrisponde uno e un solo punto);
- b) si può quindi pensare a una funzione 'conversione di un vettore nelle sue coordinate', che è ovviamente biiettiva (quindi invertibile); essendo le coordinate vettori di \mathbb{R}^n , esiste una funzione biiettiva tra qualsiasi spazio vettoriale di dimensione n e \mathbb{R}^n ;²
- c) **conseguenza importante:** è indifferente operare su un vettore o sul vettore delle sue coordinate. Ad esempio, rispetto alla base $\{(2, 0), (0, 2)\}$ le coordinate di $(4, 8)$ sono $(2, 4)$, quelle di $(6, 6)$ sono $(3, 3)$. Se sommo i vettori ottengo:

$$(4, 8) + (6, 6) = (10, 14)$$

se sommo le coordinate, ottengo le coordinate della somma dei vettori:

$$(2, 4) + (3, 3) = (5, 7), \quad 5(2, 0) + 7(0, 2) = (10, 14)$$

Analogamente, se multiplico $(4, 8)$ per lo scalare 0.5 ottengo:

$$0.5(4, 8) = (2, 4)$$

se multiplico le coordinate, ottengo le coordinate dello stesso vettore:

$$0.5(2, 4) = (1, 2), \quad 1(2, 0) + 2(0, 2) = (2, 4)$$

²Si dice che qualsiasi spazio vettoriale di dimensione n è *isomorfo* a \mathbb{R}^n , nonché, per la proprietà transitiva, che gli spazi vettoriali di uguale dimensione sono isomorfi tra loro.

Meglio ripetere: è indifferente operare sui vettori di un qualsiasi spazio vettoriale o sui vettori delle loro coordinate; unica accortezza: *se si opera sulle coordinate, ci si deve ricordare di convertire le coordinate ottenute in un vettore*. $(10, 14)$ e $(5, 7)$ si somigliano, ma si deve stare attenti a ricordare che il $(5, 7)$ ottenuto come sopra è un vettore di coordinate, il vettore delle coordinate di $(10, 14)$. Ne riparleremo.

1.5 Somma e intersezione di sottospazi

Gli spazi vettoriali sono comunque insiemi, quindi ci si può chiedere che succede se proviamo l'unione e l'intersezione di due (o più) spazi vettoriali. Meglio: dal momento che non avrebbe senso sommare polinomi e matrici, o matrici e funzioni, ci si può chiedere cosa succede se, *dato uno spazio vettoriale*, proviamo l'unione e l'intersezione di due suoi sottospazi.

Sorpresa: l'intersezione di due sottospazi è ancora uno spazio vettoriale, ma l'unione molto spesso no. Si tratterebbe di capire perché.

Immaginiamo di avere uno spazio vettoriale U e due suoi sottospazi, V e W , e scegliamo n vettori di V , $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Li scegliamo in modo tale che, se l'intersezione tra V e W non è vuota, quegli n vettori appartengono anche a W . Abbiamo visto che V è uno spazio vettoriale se qualsiasi combinazione lineare di vettori appartenenti a V è ancora un vettore appartenente a V . Ma questo vale, ovviamente, anche per W ; anche W , quindi, contiene tutte le combinazioni lineari di quei vettori. Ne segue che se $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ appartengono a $V \cap W$, gli appartengono anche tutte le loro combinazioni lineari (perché appartengono sia a V che a W), quindi anche $V \cap W$ è uno spazio vettoriale.

Ben diverso il caso dell'unione. Abbiamo i soliti due sottospazi e due vettori, $\mathbf{v} \in V$ e $\mathbf{w} \in W$, scelti in modo che non appartengano a entrambi. Bene: l'unione di V e W contiene ovviamente i due vettori, ma non la loro somma $\mathbf{v} + \mathbf{w}$, quindi non è uno spazio vettoriale (lo è solo se uno dei due sottospazi è contenuto nell'altro, ad esempio se $V \subset W \subset U$). Si preferisce quindi un'altra situazione, che evidentemente risolve il problema:

Somma di sottospazi. *Dati uno spazio vettoriale U e due suoi sottospazi V e W , la somma $V + W$ è l'insieme di tutti i vettori del tipo $\mathbf{v} + \mathbf{w}$.*

Somma diretta di sottospazi. *La somma di sottospazi la cui intersezione sia vuota, che si indica con \oplus invece che con $+$.*

Trovare la somma di due sottospazi e la sua dimensione è facile. Un sottospazio viene spesso definito o mediante un insieme di generatori, e basta vedere quali sono quelli linearmente indipendenti per trovare una base, oppure mediante un sistema di equazioni, e basta risolvere il sistema per trovare una base. Trovate due basi, basta metterle insieme ed escludere eventuali vettori linearmente dipendenti. In genere, esercizi di questo tipo non spaventano nessuno.

Trovare l'intersezione di due sottospazi e la sua dimensione risulta invece spesso più ostico. In linea di massima, il problema si risolve facilmente se i due sottospazi sono definiti mediante sistemi di equazioni: si mettono insieme i due sistemi in un sistemone (i vettori dell'intersezione devono soddisfare le condizioni sia del primo che del secondo sistema) e si trova una base.

Aiuta molto la:

Formula di Grassmann. Dato uno spazio vettoriale U e due suoi sottospazi V e W , si ha:

$$\dim(V) + \dim(W) = \dim(V + W) + \dim(V \cap W)$$

Esempio 1.11. Dati i sottospazi di \mathbb{R}^3 $V = \{(x, y, z) : x + y - z = 0\}$ e $W = \{(x, y, z) : x - y = 0\}$, si trova che una base di V è $B_V = \{(-1, 1, 0), (1, 0, 1)\}$, mentre una base di W è $B_W = \{(1, 1, 0), (0, 0, 1)\}$. Questi calcoli in genere si fanno fare. Per trovare una base di $V + W$ basta mettere insieme i quattro vettori delle due basi trovate e vedere quali sono quelli linearmente indipendenti, e anche questo in genere si sa fare. A volte si dimentica, però, come fare per trovare una base di $V \cap W$: in realtà basta trovare una base per un sottospazio che rispetti sia la definizione di V che quella di W , si tratta cioè di risolvere il sistema:

$$\begin{cases} x + y - z = 0 \\ x - y = 0 \end{cases}$$

La formula di Grassmann aiuta poi a capire se si è sbagliato qualcosa.

2 Applicazioni lineari e matrici associate

2.1 Applicazioni lineari

Un'applicazione, in generale, è una 'legge' che associa a *ciascun* elemento di un insieme, detto dominio, *uno e un solo* elemento di un altro, detto codominio (e non necessariamente distinto dal primo). Due definizioni:

Immagine di un elemento del dominio. Se f è un'applicazione e x è un elemento del dominio, l'elemento del codominio ad esso associato, $f(x)$, viene detto immagine di x . Si dice anche che, se $y = f(x)$, x è la controimmagine di y (della sua immagine).

Immagine dell'applicazione. Se D è il dominio di f , l'insieme di tutte le immagini viene detto immagine di f e si indica con $\text{Im}(f)$, o anche con $f(D)$.

Notare la differenza: nel primo caso si ha un elemento del codominio che è immagine di un elemento del dominio, nel secondo si ha l'insieme delle immagini che viene detto immagine dell'applicazione.

Si definiscono anche, e spesso, applicazioni di uno spazio vettoriale in un altro. Tra queste, rivestono particolare interesse le applicazioni *lineari*.

La linearità non è una cosa semplice, non perché sia complicata, ma perché un'applicazione, per essere lineare, deve rispettare due proprietà:

- a) *additività*: un'applicazione si dice additiva se $f(x + y) = f(x) + f(y)$;
- b) *omogeneità*: un'applicazione si dice omogenea se $f(kx) = kf(x)$.³

Ecco quindi la definizione di applicazione lineare tra spazi vettoriali:

Applicazione lineare. Un'applicazione $T : V \rightarrow W$ si dice lineare se, comunque scelti due vettori $\mathbf{v}_1, \mathbf{v}_2$ di V e due scalari h, k :

³A rigore, questa è l'omogeneità di primo grado.

a) l'immagine della somma è la somma delle immagini (additività):

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$$

b) l'immagine del prodotto per uno scalare è il prodotto per quello scalare dell'immagine (omogeneità):

$$T(h\mathbf{v}_1) = hT(\mathbf{v}_1)$$

In generale, quindi, se:

$$T(h\mathbf{v}_1 + k\mathbf{v}_2) = hT(\mathbf{v}_1) + kT(\mathbf{v}_2)$$

Problema: come capire se un'applicazione è lineare?

La risposta è ovvia: verificare che sia additiva e omogenea, e non è difficile. Soprattutto, c'è un criterio maledettamente semplice. Cominciamo comunque con un paio di esempi.

Esempio 2.1. L'applicazione

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y^2 \end{bmatrix}$$

non è additiva; infatti:

$$T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ (y_1 + y_2)^2 \end{bmatrix}$$

è diverso da:

$$T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ y_1^2 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2^2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1^2 + y_2^2 \end{bmatrix}$$

per il semplice motivo che $(y_1 + y_2)^2 \neq y_1^2 + y_2^2$. Non è nemmeno omogenea:

$$kT\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = k\begin{bmatrix} x \\ y^2 \end{bmatrix} = \begin{bmatrix} kx \\ ky^2 \end{bmatrix} \neq T\left(\begin{bmatrix} kx \\ ky \end{bmatrix}\right) = \begin{bmatrix} ky \\ k^2y^2 \end{bmatrix}$$

Si vede chiaramente che quell'elevazione al quadrato crea problemi.

Esempio 2.2. L'applicazione $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y + 1 \end{bmatrix}$ non è additiva; infatti:

$$T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 + 1 \end{bmatrix}$$

è diverso da:

$$T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ y_1 + 1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 + 1 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 + 2 \end{bmatrix}$$

Possiamo fermarci qui (se non è additiva non è lineare) e notare che questa volta è la presenza di quel +1 a creare problemi.

Conclusione: un'applicazione è lineare solo se trasforma in espressioni con termini tutti di primo grado: niente potenze con esponenti diversi da 1, niente 'termini noti'. Nel caso di spazi \mathbb{R}^n , un'applicazione è lineare solo se è del tipo:

$$T \left(\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \right) = \begin{bmatrix} p_1(x_1, \dots, x_n) \\ p_2(x_1, \dots, x_n) \\ \dots \\ p_n(x_1, \dots, x_n) \end{bmatrix}$$

dove i p_i sono polinomi nulli oppure omogenei di grado 1.

2.2 Nucleo e immagine di un'applicazione lineare (1)

Come qualsiasi funzione, un'applicazione lineare può essere iniettiva, suriettiva, biiettiva. Ricordiamo che una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ è:

- a) *iniettiva*: se ciascun elemento dell'immagine di f ha una sola controimmagine; ad esempio, $f(x) = x^2$ non è iniettiva, in quanto $f(3) = f(-3) = 9$, cioè 9, essendo immagine sia di 3 che di -3 , ha due controimmagini;
- b) *suriettiva*: se l'immagine di f coincide col codominio; ad esempio, $f(x)^2$ non è suriettiva in quanto non assume mai valori negativi (rimane fuori metà del codominio);
- c) *biiettiva*: se è sia iniettiva che suriettiva; questo vuol dire che ciascun elemento del dominio ha una sola immagine nel codominio e che ciascun elemento del codominio ha una e una sola controimmagine nel codominio (in pratica, si stabilisce una corrispondenza uno-a-uno tra tutti gli elementi del dominio e tutti quelli del codominio); ad esempio, $f(x) = 2x$ è iniettiva e suriettiva, quindi biiettiva.

Una funzione biiettiva è anche invertibile. Ad esempio, se $y = f(x) = 2x$, allora $x = f^{-1}(y) = 1/2y$.

È ovviamente importante riuscire a capire se un'applicazione lineare $T : V \rightarrow W$ è iniettiva, suriettiva, biiettiva (quindi invertibile) o nessuna delle tre. A questo scopo si introduce il concetto di *nucleo*.

Nucleo (o kernel) di un'applicazione lineare. È l'insieme dei vettori del dominio la cui immagine è il vettore nullo e si indica con $\text{Ker}(T)$:

$$\text{Ker}(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}_W\}$$

Si dimostra che il nucleo di un'applicazione lineare è un sottospazio del dominio e che l'immagine di un'applicazione lineare è un sottospazio del codominio.

Torneremo più avanti sui modi per determinare la dimensione del nucleo e dell'immagine. Per ora notiamo solo che dire che il nucleo ha dimensione 0 vuol dire che nel nucleo c'è il solo vettore nullo.

Ai fini pratici, è anche importante ricordare che si dimostra la seguente relazione tra le dimensioni del dominio, del nucleo e dell'immagine:

$$\dim V = \dim \text{Ker}(T) + \dim \text{Im}(T)$$

Soprattutto, si deve tenere a mente un altro risultato:

Teorema. *Un'applicazione lineare $T : V \rightarrow W$ è:*

- *iniettiva: se e solo se il nucleo ha dimensione 0;*
- *suriettiva: se e solo se $\text{Im}(T) = W$.*

Il secondo punto è scontato, perché si basa sulla definizione di applicazione suriettiva. Il primo è meno intuitivo.

Cominciamo dalla parte facile: se un'applicazione è iniettiva, il nucleo può contenere solo il vettore nullo (quindi ha dimensione 0) perché il vettore nullo del codominio, come qualsiasi altro, ha una sola controimmagine, e questa non può essere che il vettore nullo del dominio.

Per capire perché se un'applicazione ha nucleo di dimensione 0 allora è iniettiva, invece, si deve ragionare un po'.

Ricordiamo che, se T è iniettiva, allora ciascun elemento dell'immagine di T ha una sola controimmagine, quindi si può avere $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ solo se $\mathbf{v}_1 = \mathbf{v}_2$ ($f(x) = 2x$ è iniettiva, quindi $f(x_1) = f(x_2) = 4$ vuol dire che $x_1 = x_2 = 2$).

Assumiamo ora che il nucleo abbia dimensione 0, cioè che contenga il solo vettore nullo, e prendiamo due vettori qualsiasi di V , \mathbf{v}_1 e \mathbf{v}_2 , tali che $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ (se pensiamo che l'applicazione non è iniettiva, dobbiamo poterli trovare). Si ha ovviamente: $T(\mathbf{v}_1) - T(\mathbf{v}_2) = \mathbf{0}_W$. Ma T è lineare, in particolare additiva, e questo vuol dire che: $T(\mathbf{v}_1) - T(\mathbf{v}_2) = T(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}_W$. Se il nucleo ha dimensione 0, allora deve necessariamente essere: $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}_V$, ovvero $\mathbf{v}_1 = \mathbf{v}_2$. Quindi da $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ segue $\mathbf{v}_1 = \mathbf{v}_2$, e questo vuol dire che l'applicazione è iniettiva.

2.3 Matrici associate ad applicazioni lineari

Tasto dolente. La definizione è relativamente semplice:

Matrice associata a un'applicazione lineare. *Data un'applicazione lineare $T : V \rightarrow W$, fissate le basi $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ per V e $C = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m\}$ per W (quindi V ha dimensione n e W ha dimensione m) la matrice associata all'applicazione è una matrice che ha m righe e n colonne la cui j -esima colonna, $j = 1, 2, \dots, n$, è costituita dalle coordinate rispetto a C dell'immagine del j -esimo elemento di B .*

Si nota spesso, tuttavia, che sfugge un elemento fondamentale: *le applicazioni operano sui vettori di uno spazio vettoriale, le matrici operano su (vettori di) coordinate rispetto a basi fissate.*

Dovrebbe essere evidente (se ho un'applicazione il cui dominio è uno spazio vettoriale di polinomi, come faccio a moltiplicare una matrice per un polinomio?), ma purtroppo molti esercizi riguardano applicazioni $\mathbb{R}^n \rightarrow \mathbb{R}^m$. Quindi ci si abitua a ragionare solo in termini di n -uple di numeri reali e poi, se capitano spazi di matrici o di polinomi... si scrivono messaggi disperati sul forum di www.matematicamente.it.

Temo sfugga anche che, potendo essere infinite le basi di uno spazio vettoriale, a una stessa applicazione è possibile associare tante matrici quante sono le due basi che scelgo, quindi *infinite matrici*.

Infine si nota spesso che, se pure si riesce a imparare a memoria la definizione, se pure si riesce a svolgere qualche esercizio, non si capisce assolutamente perché

una matrice associata a un'applicazione lineare debba essere di quel tipo. E questo certo non aiuta a evitare pasticci.

Cominciamo da qui.

Ho un'applicazione lineare $T : V \rightarrow W$ e voglio trovare una matrice associata. Voglio cioè passare da $T(\mathbf{v}) = \mathbf{w}$ a $\mathbf{Ax} = \mathbf{y}$.

Sottolineo che \mathbf{x} è il vettore delle coordinate di \mathbf{v} rispetto alla base di V , \mathbf{y} è il vettore delle coordinate di \mathbf{w} rispetto alla base di W , ma su questo torneremo.

Ora: \mathbf{Ax} non è altro che una combinazione lineare delle colonne di \mathbf{A} i cui coefficienti sono i componenti di \mathbf{x} . Infatti:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

Chiarito questo, proviamo a sviluppare $T(\mathbf{v}) = \mathbf{w}$.

Essendo \mathbf{v} un vettore di V , può essere espresso come combinazione lineare degli elementi della base B di V :

$$T(\mathbf{v}) = T(k_1\mathbf{b}_1 + k_2\mathbf{b}_2 + \dots + k_n\mathbf{b}_n)$$

Dal momento, però, che T è lineare, posso scrivere:

$$T(\mathbf{v}) = k_1T(\mathbf{b}_1) + k_2T(\mathbf{b}_2) + \dots + k_nT(\mathbf{b}_n)$$

Siamo arrivati a:

$$\begin{bmatrix} T(\mathbf{b}_1) & T(\mathbf{b}_2) & \dots & T(\mathbf{b}_n) \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix} = k_1T(\mathbf{b}_1) + k_2T(\mathbf{b}_2) + \dots + k_nT(\mathbf{b}_n)$$

ci siamo cioè avvicinati a qualcosa (il primo fattore del primo membro) che assomiglia ad una matrice.⁴

I vari $T(\mathbf{b}_j)$ sono però immagini di vettori, quindi vettori del codominio, e possono quindi essere espressi come combinazioni lineari degli elementi della base C di W . Basta sostituirli con i vettori colonna delle loro coordinate rispetto a C e si ottiene la matrice associata (rispetto a B e C).

Chi vuole può provare un approfondimento (consigliato). Che vuol dire «basta sostituirli ecc.»? Riscrivo i vari $T(\mathbf{b}_j)$ come combinazioni lineari:

$$T(\mathbf{b}_j) = a_{1j}\mathbf{c}_1 + a_{2j}\mathbf{c}_2 + \dots + a_{mj}\mathbf{c}_m$$

e riscrivo il secondo membro così:

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_m \end{bmatrix} \begin{bmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{bmatrix} = a_{1j}\mathbf{c}_1 + a_{2j}\mathbf{c}_2 + \dots + a_{mj}\mathbf{c}_m$$

dove il primo fattore del primo membro è un vettore di vettori, quindi una matrice che ha per colonne gli elementi della base C .

⁴In realtà lo è, perché i $T(\mathbf{b}_j)$ sono vettori.

‘Tralascio’ questa matrice e mi rimane il vettore colonna dei coefficienti della combinazione lineare, quindi:

$$\begin{bmatrix} T(\mathbf{b}_1) & T(\mathbf{b}_2) & \dots & T(\mathbf{b}_j) & \dots & T(\mathbf{b}_n) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

Ecco la nostra matrice!

Ma quel ‘tralasciare’ ha un prezzo: per ottenere davvero l’immagine \mathbf{w} di \mathbf{v} , devo moltiplicare quello che ottengo per la matrice degli elementi di C ; questo vuol dire che \mathbf{Ax} non mi dà un elemento di W , ma solo le sue coordinate rispetto alla base C !

Per ottenere \mathbf{w} devo quindi moltiplicare $\mathbf{Ax} = \mathbf{y}$ per la matrice dei vettori di C che avevo ‘tralasciato’:

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_m \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{bmatrix} = y_1 \mathbf{c}_1 + y_2 \mathbf{c}_2 + \dots + y_m \mathbf{c}_m = \mathbf{w}$$

Seguono un paio di esempi.

2.3.1 Esempi di matrici associate (1)

Ecco un esempio tanto facile da risultare ingannevole.

Esempio 2.3. Data l’applicazione $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x-y \end{bmatrix}$, trovare la matrice associata rispetto alla base canonica, $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$. Usare la matrice per trovare l’immagine del vettore $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

Cerco le immagini degli elementi della base:

$$\begin{aligned} T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

Metto in colonna e ho la matrice: $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Moltiplico ora la matrice per il vettore dato:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$$

È giusto? Sì. Solo che è troppo facile e troppi esercizi di questo tipo mettono strane idee in testa.

Esempio 2.4. Data l'applicazione $T : S_2^+ \rightarrow \mathbb{R}_3[t]$

$$T\left(\begin{bmatrix} a & b \\ b & c \end{bmatrix}\right) = a - b + (b - c)t + (b - c)t^2 + (a - b)t^3$$

un'applicazione dallo spazio delle matrici simmetriche di ordine 2 nello spazio dei polinomi in una variabile di grado non superiore a 3, trovare la matrice associata rispetto alle basi:

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}, \quad C = \{1, t, t^2, t^3\}$$

e usare la matrice per trovare l'immagine della matrice $\begin{bmatrix} 1 & 5 \\ 5 & 3 \end{bmatrix}$.

Cerco le immagini degli elementi della base:

$$\begin{aligned} T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) &= 1 + t^3 \\ T\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) &= -1 + t + t^2 - t^3 \\ T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) &= -t - t^2 \end{aligned}$$

E che ci faccio? Posso forse usarli per costruire una matrice? Costruisco forse una matrice di tre righe e una colonna i cui elementi siano... polinomi?

Ovviamente no. Devo prima convertire quei polinomi in vettori di coordinate rispetto alla base C :

$$\begin{aligned} \text{Coord}(1 + t^3) &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \\ \text{Coord}(-1 + t + t^2 - t^3) &= \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \\ \text{Coord}(-t - t^2) &= \begin{bmatrix} 0 \\ -1 \\ -1 \\ 0 \end{bmatrix} \end{aligned}$$

A questo punto costruire la matrice associata all'applicazione è uno scherzo:

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

Ora si tratta di usare la matrice. Posso forse moltiplicarla per la matrice data? Ovviamente no: non posso moltiplicare una matrice 4×3 per una 2×2 . Posso solo (e *devo*) moltiplicare la matrice \mathbf{A} per il vettore delle coordinate, rispetto alla base B , della matrice data:

$$\text{Coord} \left(\begin{bmatrix} 1 & 5 \\ 5 & 3 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \\ 2 \\ -4 \end{bmatrix}$$

Ho forse finito? **NO!**

Quello che ho ottenuto è un vettore di \mathbb{R}^4 , non è mica un polinomio! L'immagine della matrice data deve essere un polinomio, ed è facilissimo individuarlo:

$$-4 + 2t + 2t^2 - 4t^3$$

Ma troppe volte si dimentica quest'ultimo passaggio...

Si deve ricordare che spazi come \mathbb{R}^n con base canonica costituiscono un'eccezione: il solo caso in cui un vettore e il vettore delle sue coordinate coincidono.

Si deve ricordare che, mentre le applicazioni operano su vettori, le matrici associate operano su vettori di coordinate e restituiscono vettori di coordinate.

Si deve ricordare che in generale, escluso il caso di spazi come \mathbb{R}^n con base canonica, si devono convertire in vettori di coordinate le immagini degli elementi della base di partenza e si deve poi convertire il vettore di coordinate ottenuto in un elemento dello spazio di arrivo.

2.3.2 Esempi di matrici associate (2)

Negli esempi precedenti erano date sia l'applicazione in forma generale che le basi (per 'forma generale' intendo che viene data l'immagine del generico vettore del dominio). Altre volte vengono date solo le immagini di alcuni vettori e si deve risalire all'applicazione.

Osservazione. Se un'applicazione lineare $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ viene data in forma generale, la matrice \mathbf{A} associata *rispetto alle basi canoniche* si trova subito, senza bisogno di calcoli; basta scrivere la matrice che ha in ciascuna riga elementi uguali ai coefficienti degli elementi del generico vettore del dominio che intervengono nella definizione del corrispondente elemento dell'immagine:

$$T \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} -x + y + z \\ y - z \\ x + z \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

Infatti:

$$\begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x + y + z \\ y - z \\ x + z \end{bmatrix}$$

Da notare che \mathbf{A} è la matrice associata *rispetto alle basi canoniche* in quanto si è ragionato in termini di vettori del dominio e del codominio, e *solo rispetto alle basi canoniche vettori e vettori di coordinate coincidono*.

Viceversa, se si ha la matrice associata all'applicazione rispetto alle basi canoniche del dominio e del codominio, si risale immediatamente all'applicazione.

Questo vuol dire che, se si hanno solo le immagini di alcuni vettori, per risalire da essi all'applicazione si deve cercare la matrice associata alle basi canoniche.

Esempio 2.5. Sia $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ un'applicazione lineare tale che:

$$T\left(\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad T\left(\begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \quad T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Per trovare l'applicazione, mi serve la matrice associata rispetto alle basi canoniche. Per trovare questa, occorrono le immagini degli elementi della base canonica, ma ne ho solo una, quella di $(0, 1, 0)$. Che faccio?

Nei casi più complessi può essere necessario impostare un piccolo sistema, qualcosa del tipo:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = a \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} a = 1 \\ -2a - 2b + c = 0 \\ b = 0 \end{cases}$$

si tratta cioè di esprimere ciascun vettore della base canonica come combinazione lineare dei vettori del dominio che conosco e trovare i relativi coefficienti. Come si vede, in questo caso il sistema è di facilissima soluzione e possiamo saltare i passaggi (ci si arriva anche 'a occhio'):

$$\begin{aligned} T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) &= T\left(\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}\right) + 2T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) &= T\left(\begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}\right) + 2T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -3 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

Mi basta mettere in colonna le immagini per ottenere la matrice:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

da cui l'applicazione:

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + z \\ y - z \end{bmatrix}$$

2.4 Nucleo e immagine di un'applicazione lineare (2)

Abbiamo già visto, nella sezione 2.2, cosa sono il nucleo e l'immagine di un'applicazione lineare e, in particolare, a cosa serve il nucleo (a capire se l'applicazione è iniettiva). Ora vediamo come si trovano, grazie a una matrice associata, il nucleo e l'immagine.

2.4.1 Trovare il nucleo

Trovare il nucleo è facile. Data un'applicazione $T : V \rightarrow W$, con $\dim V = n$ e $\dim W = m$, se una volta fissate le basi si è trovata la matrice $m \times n$ associata \mathbf{A} , la definizione del nucleo:

$$\text{Ker}(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}_W\}$$

diventa:

$$\text{Ker}(T) = \{\mathbf{v} \in V : \mathbf{x} = \text{Coord}(\mathbf{v}), \mathbf{Ax} = \mathbf{0}_{\mathbb{R}^m}\}$$

(ricordo che una matrice opera *sempre* su vettori di coordinate).

Si tratta quindi di risolvere un normale sistema lineare omogeneo $\mathbf{Ax} = \mathbf{0}$ per trovare una base (quindi anche la dimensione) del nucleo.

Esempio 2.6. Riprendiamo l'esempio 2.4 già proposto: un'applicazione

$$T\left(\begin{bmatrix} a & b \\ b & c \end{bmatrix}\right) = a - b + (b - c)t + (b - c)t^2 + (a - b)t^3$$

con matrice associata

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

rispetto alle basi

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}, \quad C = \{1, t, t^2, t^3\}$$

Costruiamo il sistema $\mathbf{Ax} = \mathbf{0}$ e risolviamolo. Riducendo la matrice a gradini otteniamo un sistema equivalente più semplice e il relativo insieme delle soluzioni (queste cose in genere si sanno fare):

$$\mathbf{Ax} = \mathbf{0} \Rightarrow \begin{cases} x_1 - x_3 = 0 \\ x_2 - x_3 = 0 \end{cases} \Rightarrow \text{Sol}(\mathbf{A}, \mathbf{0}) = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}$$

La dimensione del nucleo è chiaramente 1, ma...

Attenzione! Quel vettore $(1, 1, 1)$ *non* è l'elemento di una base del nucleo!

Lo sarebbe se il dominio fosse \mathbb{R}^3 (il nucleo è un sottospazio del dominio) e se la matrice fosse associata all'applicazione rispetto alle basi canoniche. Qui il dominio è uno spazio vettoriale di *matrici simmetriche di ordine 2*, quindi bisogna passare da quello che si è ottenuto, un *vettore di coordinate*, al corrispondente vettore del dominio, cioè alla corrispondente matrice. Essendo $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

la base fissata per il dominio, una base del nucleo ha come unico elemento la matrice $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Infatti:

$$T\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) = 1 - 1 + (1 - 1)t + (1 - 1)t^2 + (1 - 1)t^3 = 0$$

2.4.2 Trovare l'immagine

Trovare l'immagine è altrettanto facile, ma si rischia di inciampare. . . Analogamente a quanto abbiamo già visto per il nucleo, data un'applicazione $T : V \rightarrow W$, con $\dim V = n$ e $\dim W = m$, se una volta fissate le basi si è trovata la matrice $m \times n$ associata A , la definizione dell'immagine:

$$\text{Im}(T) = \{\mathbf{w} \in W : \mathbf{w} = T(\mathbf{v}), \mathbf{v} \in V\}$$

diventa:

$$\text{Im}(T) = \{\mathbf{w} \in W : \mathbf{w} = \text{Coord}^{-1}(\mathbf{y}), \mathbf{A}\mathbf{x} = \mathbf{y}, \mathbf{x} = \text{Coord}(\mathbf{v}), \mathbf{v} \in V\}$$

ovvero: l'immagine è costituita da quelli elementi di W le cui coordinate rispetto alla base fissata per il codominio sono il risultato del prodotto di \mathbf{A} per le coordinate di un qualsiasi elemento del dominio.

In pratica, *purché si ricordi che le matrici operano solo su coordinate*, basta limitarsi a $\mathbf{A}\mathbf{x} = \mathbf{y}$ ricordando che il prodotto $\mathbf{A}\mathbf{x}$ non è altro che una combinazione lineare delle colonne di \mathbf{A} i cui coefficienti sono i componenti di \mathbf{x} (cfr. sezione 2.3).

Questo vuol dire che l'immagine di T , in quanto generata dal prodotto $\mathbf{A}\mathbf{x}$,⁵ ha una dimensione pari al numero di colonne di \mathbf{A} linearmente indipendenti.

Conseguenza: la dimensione dell'immagine non è altro che il rango di \mathbf{A} .

Possibile impiccio: per trovare una base dell'immagine si devono trovare le *colonne* linearmente indipendenti di \mathbf{A} , ma si è troppo abituati a ridurre una matrice per righe. . .

Soluzioni:

- a) si può ridurre la matrice \mathbf{A} per colonne, ma è meglio evitare perché, essendo abituati a ridurre per righe, si rischiano errori tanto banali quanto probabili;
- b) si riduce la matrice per righe e poi si prendono le colonne di \mathbf{A} corrispondenti ai pivot della matrice ridotta; da ricordare che *non si possono prendere le colonne della ridotta*;
- c) si traspone la matrice (facile) e si riduce per righe la trasposta; in questo caso, si possono prendere tranquillamente le righe non nulle della ridotta che, rimesse in colonna, costituiscono le coordinate degli elementi di una base dell'immagine anche se non sono uguali alle colonne di \mathbf{A} .

Esempio 2.7. Proseguiamo l'esempio precedente. Ridurre \mathbf{A} per righe si può fare in tanti modi; diciamo che una possibile riduzione potrebbe essere:

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

⁵Qualcuno potrebbe obiettare: ma come? Non hai detto che le matrici operano solo su coordinate? E allora come fai a dire che l'immagine è generata da un prodotto per un vettore di coordinate? Risposta: come avevo mostrato nella sezione 1.4, è *indifferente operare sui vettori di un qualsiasi spazio vettoriale o sui vettori delle loro coordinate, a condizione che si ricordi che se si opera sulle coordinate si devono poi convertire le coordinate in vettori* (questo perché qualsiasi spazio vettoriale di dimensione n è isomorfo a \mathbb{R}^n , che è lo spazio delle coordinate dei suoi vettori rispetto a una base).

Nella ridotta ci sono due pivot (la radice ha rango 2), prendo le corrispondenti colonne di A e ho trovato una base:

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \right\}$$

Giusto? **Sbagliato!** Ho ottenuto due vettori di *coordinate* degli elementi di una base. L'immagine è un sottospazio di polinomi, quindi la sua base così trovata è:

$$\{1 + t^3, -1 + t + t^2 - t^3\}$$

Da notare che, *se avessi preso le colonne con i pivot della ridotta, non sarei arrivato a una base*. Sarei infatti arrivato a $\{1, t\}$, cioè a due polinomi che non possono generare polinomi di grado superiore a 1, ma avevamo già visto nell'esempio 2.4 che, per quanto l'immagine sia un sottospazio del codominio, l'applicazione può convertire matrici simmetriche anche in polinomi di terzo grado.

Proviamo ora a trasporre e a ridurre per righe:

$$\mathbf{A}^T = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 1 & -1 \\ 0 & -1 & -1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Prendo le righe linearmente indipendenti, le rimetto in colonna e ottengo... le coordinate degli elementi di una base:

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

cioè la base trovata è:

$$\{1 + t^3, t + t^2\}$$

Il primo vettore di coordinate è uguale alla prima colonna di \mathbf{A} , il secondo no, ma è uguale alla terza moltiplicata per -1 ; notiamo però che, così come i polinomi della base trovata prima possono generare quelli della base trovata ora, anche questi possono generare quelli. Infatti, il primo polinomio è comune e, per il resto:

$$\begin{aligned} -1 + t + t^2 - t^3 &= -(1 + t^3) + (t + t^2) \\ t + t^2 &= (-1 + t + t^2 - t^3) + (1 + t^3) \end{aligned}$$

quindi entrambe le basi vanno bene.

2.5 Teorema della nullità e del rango

Argomento breve, ma importante.

Avevamo visto nella sezione 2.2 che:

- a) data un'applicazione lineare $T : V \rightarrow W$, con V di dimensione n e W di dimensione m , si ha:

$$n = \dim V = \dim \text{Ker}(T) + \dim \text{Im}(T)$$

- b) un'applicazione lineare T è iniettiva se e solo se $\dim \text{Ker}(T) = 0$.

Avevamo poi visto, parlando di matrici associate, che:

- c) tali matrici hanno un numero di colonne n pari alla dimensione del dominio e un numero di righe m pari alla dimensione del codominio;
 d) la dimensione dell'immagine di un'applicazione lineare T è pari al numero di colonne linearmente indipendenti della matrice associata \mathbf{A} , cioè al suo rango.

Infine, avevamo visto nella sezione 2.4.1 che:

- e) il nucleo di un'applicazione lineare T , cui sia associata la matrice \mathbf{A} , si trova risolvendo un sistema lineare omogeneo $\mathbf{Ax} = \mathbf{0}$.

Chiamiamo *nucleo* di una matrice \mathbf{A} lo spazio delle soluzioni del sistema $\mathbf{Ax} = \mathbf{0}$ e *nullità* la sua dimensione, che corrisponde in tutto e per tutto alla dimensione del nucleo, o kernel, dell'applicazione cui la matrice è associata.

Sostituendo $\dim V$ con n , $\dim \text{Ker}(T)$ con $\text{null}(\mathbf{A})$ e $\dim \text{Im}(T)$ con $\text{rk}(\mathbf{A})$, otteniamo:

$$n = \text{null}(A) + \text{rk}(A)$$

Ovvero: il numero di colonne di una matrice è pari alla somma del suo rango e della sua nullità.

Questo consente di indagare le caratteristiche di un'applicazione lineare osservando una sua matrice associata. In particolare:

- una volta trovato il rango r della matrice, la dimensione del nucleo dell'applicazione è $n - r$;
- se una matrice ha rango uguale al numero delle colonne (quindi le sue colonne sono linearmente indipendenti), allora ha nullità 0, quindi l'applicazione è iniettiva;
- quando dominio e codominio hanno la stessa dimensione, dunque quando la matrice associata è quadrata, se questa ha rango pieno allora è invertibile, quindi è tale anche l'applicazione (esiste l'applicazione inversa). È infatti anche suriettiva, perché la dimensione dell'immagine è uguale a quella del codominio.

2.6 Matrici di cambiamento di base

Una matrice di cambiamento di base è una matrice associata a un'applicazione lineare di uno spazio vettoriale in se stesso,⁶ essendo quell'applicazione nient'altro che l'applicazione identità, che associa ciascun vettore a se stesso (ma rispetto a una base diversa).

Invece di avere $T : V \rightarrow W$, con B base di V e C base di W , si ha, più semplicemente, $Id : V \rightarrow V$, con B base di partenza e C base di arrivo.

Invece di calcolare l'immagine di ogni elemento della base di partenza, per convertirla nel vettore di coordinate rispetto alla base di arrivo, si convertono direttamente gli elementi della base di partenza, in quanto $T(\mathbf{v}) = \mathbf{v}$. In sostanza, se si ha

⁶Un'applicazione lineare di uno spazio vettoriale in se stesso viene anche detta *operatore lineare*.

familiarità con le matrici associate ad applicazioni lineari, si sa già tutto quello che occorre sapere.

Si possono tuttavia studiare le matrici di cambiamento di base anche prima del caso più generale, usando una definizione come la seguente:

Matrice di cambiamento di base. *Dati uno spazio vettoriale V di dimensione n e due sue basi B e C , la matrice del cambiamento di base da B a C è una matrice quadrata di ordine n le cui colonne sono costituite dalle coordinate degli elementi della base di partenza, B , rispetto alla base d'arrivo C .*

Per capire come mai la matrice sia fatta in quel modo, si può pensare a un semplice esempio in cui V abbia dimensione 2, con $B = \{\mathbf{b}_1, \mathbf{b}_2\}$ e $C = \{\mathbf{c}_1, \mathbf{c}_2\}$.

Scelto un elemento \mathbf{v} di V , lo esprimiamo come combinazione lineare degli elementi di B :

$$\mathbf{v} = x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2$$

quindi (x_1, x_2) è il vettore delle sue coordinate rispetto a B .

Dato che ci interessano le coordinate di \mathbf{v} rispetto a C , esprimiamo \mathbf{b}_1 e \mathbf{b}_2 come combinazioni lineari degli elementi di C :

$$\mathbf{b}_1 = a_{11} \mathbf{c}_1 + a_{21} \mathbf{c}_2$$

$$\mathbf{b}_2 = a_{12} \mathbf{c}_1 + a_{22} \mathbf{c}_2$$

Scriviamo quindi:

$$\begin{aligned} \mathbf{v} &= x_1(a_{11} \mathbf{c}_1 + a_{21} \mathbf{c}_2) + x_2(a_{12} \mathbf{c}_1 + a_{22} \mathbf{c}_2) = \\ &= x_1 a_{11} \mathbf{c}_1 + x_2 a_{12} \mathbf{c}_1 + x_1 a_{21} \mathbf{c}_2 + x_2 a_{22} \mathbf{c}_2 = \\ &= (x_1 a_{11} + x_2 a_{12}) \mathbf{c}_1 + (x_1 a_{21} + x_2 a_{22}) \mathbf{c}_2 \end{aligned}$$

Il vettore delle coordinate di \mathbf{v} rispetto a C è quindi

$$\begin{bmatrix} x_1 a_{11} + x_2 a_{12} \\ x_1 a_{21} + x_2 a_{22} \end{bmatrix}$$

Per ottenerlo, basta moltiplicare il vettore delle coordinate rispetto a B , (x_1, x_2) , per una matrice 2×2 che abbia come colonne le coordinate degli elementi di B rispetto a C ; infatti:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 a_{11} + x_2 a_{12} \\ x_1 a_{21} + x_2 a_{22} \end{bmatrix}$$

Questo vuol dire che, se ho le coordinate rispetto a B di un qualsiasi vettore, moltiplicando quella matrice per esse ottengo le coordinate rispetto a C dello stesso vettore.

Abbiamo già notato che le matrici di cambiamento di base sono *tutte quadrate*. Infatti, se moltiplico una tale matrice per un vettore di n componenti, la matrice deve avere n colonne; ma devo ottenere ancora un vettore di n componenti, quindi la matrice deve avere anche n righe.

Aggiungiamo ora che tali matrici sono anche *sempre invertibili*. Un'applicazione identità, che manda ciascun vettore in se stesso, è infatti chiaramente biettiva. In

particolare, è iniettiva e quindi il suo nucleo ha dimensione 0. Il teorema della nullità e del rango (sez. 2.5) ci dice che la matrice quadrata associata, avendo nullità 0, ha rango pari al numero delle colonne, quindi è invertibile.

Esempio 2.8. Immaginiamo di lavorare con lo spazio vettoriale \mathbb{R}^3 e di rappresentarlo come un familiare spazio tridimensionale: se scegliamo un qualsiasi vettore, diciamo $\mathbf{v} = (3, 2, 1)$, avremo che 3 è l'ascissa, 2 l'ordinata e 1 la quota. Il normale sistema di riferimento non è altro che la base canonica: $(1, 0, 0)$ è il vettore unitario (l'unità di misura) per l'asse x delle ascisse, $(0, 1, 0)$ quello per l'asse y delle ordinate, $(0, 0, 1)$ quello per l'asse z delle quote. Usando questo sistema di riferimento (usando la base canonica E), un vettore e il vettore delle sue coordinate coincidono: $(3, 2, 1)$ è, al tempo stesso, sia un vettore (un punto) di \mathbb{R}^3 , sia il suo vettore di coordinate rispetto alla base canonica.

Ora cambiamo sistema di riferimento (cambiamo base), orientando l'asse delle ordinate nel verso opposto e scambiando l'ordinata con la quota. In sostanza, la nuova base diventa

$$E' = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \right\}$$

Facile vedere che, rispetto alla nuova base, il vettore \mathbf{v} ha coordinate $(3, 1, -2)$. Proviamo a costruire la matrice di cambiamento di base e verifichiamo che trasformi effettivamente le coordinate in questo modo.

Troviamo le coordinate rispetto a E' degli elementi di E :

$$\begin{aligned} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} &= 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} && \text{quindi: } \text{Coord} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} &= 0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 1 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} && \text{quindi: } \text{Coord} \left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} &= 0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} && \text{quindi: } \text{Coord} \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

Verifichiamo ora che la matrice che abbia per colonne le coordinate trovate trasforma effettivamente $v = (3, 2, 1)$ nel vettore di coordinate $(3, 1, -2)$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$$

Funziona.

Da notare che $(3, 1, -2)$ *non è un altro vettore! È sempre il vettore $(3, 2, 1)$, questa volta però espresso in termini delle sue coordinate rispetto alla base di arrivo E' .* Infatti,

$$3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Verifichiamo anche che la matrice, una volta invertita, ci consente di passare dalle coordinate rispetto a E' a quelle rispetto alla base canonica:

$$\text{se } \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \quad \text{allora } \mathbf{A}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

infatti:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

moltiplicando \mathbf{A}^{-1} per le coordinate di \mathbf{v} rispetto a E' :

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

otteniamo di nuovo le coordinate di \mathbf{v} rispetto alla base canonica, quindi \mathbf{v} ‘in persona’.

Domanda: come ottenere la matrice inversa? Normalmente si sa fare (o almeno si spera...), ma in casi come quello dell'esempio – non poi così rari negli esercizi – si può usare un semplice truccetto. Dato che le coordinate di un qualsiasi vettore di \mathbb{R}^n rispetto alla base canonica coincidono con i componenti del vettore, per trovare la matrice di cambiamento da una base qualsiasi alla base canonica basta mettere in colonna gli elementi di quella base. Si può notare, infatti, che la matrice \mathbf{A}^{-1} ha per colonne niente altro che gli elementi della base E' .

2.7 Matrici simili (primi cenni)

Un'applicazione identità trasforma ciascun vettore in se stesso; se la indichiamo con Id , abbiamo $Id(\mathbf{v}) = \mathbf{v}$.

Una matrice associata a un'applicazione identità rispetto a basi diverse di partenza e di arrivo, detta matrice di cambiamento di base, è sempre quadrata e invertibile.

L'applicazione identità è solo un caso particolare delle applicazioni di uno spazio in se stesso, dette *operatori lineari*.

A un operatore $T : V \rightarrow V$ si possono associare tante matrici quante sono le basi fissate per il dominio e il codominio.⁷

Se ho un operatore lineare e una matrice a esso associata rispetto a una base, come posso trovare la matrice associata rispetto a un'altra base?

Consideriamo un caso semplice: ho un'applicazione $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ con base B sia per il dominio che per il codominio e cerco la relativa matrice associata. Conosco però la matrice associata rispetto alla base canonica, magari perché l'applicazione

⁷Si intende che dominio e codominio hanno la stessa base; in generale si usa così quando si lavora su operatori lineari. Ovvio eccezione proprio l'applicazione identità: se le basi fossero uguali, la matrice associata sarebbe, banalmente, la matrice identità.

viene definita in quella che altrove (sez. 2.3.2) ho chiamato ‘forma generale’. Ad esempio, se l’applicazione è definita così:

$$T \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} x + y - z \\ y + z \\ 2x \end{bmatrix}$$

la matrice associata rispetto alla base canonica E si trova subito⁸ ed è:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 2 & 0 & 0 \end{bmatrix}$$

Immaginiamo di voler trovare la matrice associata a T rispetto alla base

$$B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

In sostanza cerco una matrice che, moltiplicata per le coordinate rispetto a B di un vettore, mi dia le coordinate rispetto a B dell’immagine di quel vettore secondo l’operatore T .

Potrei anche usare la matrice \mathbf{A} , a condizione di ‘darle in pasto’ coordinate rispetto alla base canonica e di convertire poi il risultato (espresso in coordinate rispetto alla base canonica) in coordinate rispetto a B . Avrei quindi bisogno di tre pezzi:

- a) una matrice di cambiamento di base da B a E (la base canonica), che indichiamo con \mathbf{N} ;
- b) la matrice \mathbf{A} ;
- c) una matrice di cambiamento di base da E a B , e non sarà altro che l’inversa di \mathbf{N} , quindi \mathbf{N}^{-1} .

Ho già \mathbf{A} e, per quanto visto a proposito delle matrici di cambiamento di base, trovare \mathbf{N} è semplicissimo: è infatti una matrice che ha per colonne gli elementi di B . Rimane poi da trovarne l’inversa, e si può fare in due modi:

- a) calcolo l’inversa;
- b) costruisco una matrice che abbia per colonne le coordinate rispetto a B degli elementi di E .

Arrivo comunque a:

$$\mathbf{N} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{N}^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

Proviamo.

⁸Chi non ricordasse perché potrebbe tornare un attimo alla sez. 2.3.2.

Ho il vettore $\mathbf{v} = (2, 6, 3)$, che l'operatore T trasforma in $\mathbf{w} = (5, 9, 4)$. Per ottenere \mathbf{w} si può applicare la definizione di T , ma si può anche usare la matrice \mathbf{A} ; essendo \mathbf{v} uguale al vettore delle sue coordinate rispetto alla base canonica:

$$\mathbf{w} = T(\mathbf{v}) = \mathbf{A}\mathbf{v} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 9 \\ 4 \end{bmatrix}$$

Voglio ora partire dalle coordinate di \mathbf{v} rispetto alla base B , $\mathbf{x} = (7, 4, -1)$, per ottenere le coordinate di $T(\mathbf{v})$ sempre rispetto a B . Eseguo tre passaggi:

1) converto \mathbf{x} in \mathbf{v} usando la matrice \mathbf{N} :

$$\mathbf{v} = \mathbf{N}\mathbf{x} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 3 \end{bmatrix}$$

2) eseguo $\mathbf{A}\mathbf{v}$ come sopra per ottenere l'immagine di \mathbf{v} , $\mathbf{w} = (5, 9, 4)$;

3) converto \mathbf{w} nel vettore \mathbf{y} delle sue coordinate rispetto a B , che è $(9, 4, 0)$, usando la matrice \mathbf{N}^{-1} :

$$\mathbf{y} = \mathbf{N}^{-1}\mathbf{w} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 9 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 4 \\ 0 \end{bmatrix}$$

L'aspetto interessante è che posso 'mettere in fila' i tre passaggi risalendo dal terzo al primo:

$$\mathbf{y} = \mathbf{N}^{-1}\mathbf{w} = \mathbf{N}^{-1}\mathbf{A}\mathbf{v} = \mathbf{N}^{-1}\mathbf{A}\mathbf{N}\mathbf{x}$$

Ci vuol molto a vedere che $\mathbf{N}^{-1}\mathbf{A}\mathbf{N}$ non è altro che la matrice associata a T rispetto alla base B ? Verifichiamo:

$$\mathbf{N}^{-1}\mathbf{A}\mathbf{N} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 6 & 1 \\ -1 & 3 & 1 \\ 3 & -5 & 1 \end{bmatrix} = \mathbf{M}$$

Se eseguiamo il prodotto $\mathbf{M}\mathbf{x}$ otteniamo \mathbf{y} . Provare per credere...

Due matrici \mathbf{A} e \mathbf{M} legate dalla relazione $\mathbf{M} = \mathbf{N}^{-1}\mathbf{A}\mathbf{N}$, oppure – ed è lo stesso – $\mathbf{A} = \mathbf{N}\mathbf{M}\mathbf{N}^{-1}$, vengono dette *matrici simili*. Si tratta di un argomento molto più ampio di quanto può sembrare da queste note preliminari, ma possiamo già notare che *matrici simili rappresentano uno stesso operatore lineare rispetto a basi diverse*.

3 Autovalori, autovettori, autospazi

3.1 Le definizioni e il loro senso

Le definizioni sono semplici. Un po' meno semplice è capirne il senso.

Autovalore, autovettore, autospazio. Dato un operatore lineare $T : V \rightarrow V$, un vettore non nullo \mathbf{v} di V viene detto autovettore per T se esiste uno scalare λ tale che:

$$T(\mathbf{v}) = \lambda\mathbf{v}$$

λ viene detto *autovalore relativo a \mathbf{v}* . Si dice inoltre *autospazio relativo a λ* il *sottospazio*:

$$V_\lambda = \{\mathbf{v} \in V : T(\mathbf{v}) = \lambda \mathbf{v}\}$$

Prima osservazione banale: un vettore nullo non è autovettore per il semplice motivo che si ha sempre $T(\mathbf{0}_V) = \lambda \mathbf{0}_V = \mathbf{0}_V$ quale che sia λ .

Per cercare comunque di capire il senso delle definizioni, prendiamo lo spazio vettoriale \mathbb{R}^3 , lo rappresentiamo come il normale spazio tridimensionale (quello con ascissa, ordinata e quota) e immaginiamo un paio di facili operatori.

Esempio 3.1. Sia $F(x, y, z) = (x, y, 2z)$. Prendiamo il vettore $\mathbf{v} = (0, 0, 90)$ e proviamo:

$$F(0, 0, 90) = (0, 0, 180) = 2(0, 0, 90)$$

Chiaramente \mathbf{v} è un autovettore e 2 è il relativo autovalore. Proviamo ora con $\mathbf{w} = (0, 0, 85)$:

$$F(0, 0, 85) = (0, 0, 170) = 2(0, 0, 85)$$

Anche \mathbf{w} è un autovettore. Può stupire? Non molto: F lascia invariate ascissa e ordinata e raddoppia la quota, quindi qualsiasi vettore che abbia i primi due componenti nulli verrà raddoppiato. Questo vuol dire che, una volta trovato un autovettore, se ne sono trovati infiniti (tutti quelli con i primi due componenti nulli); non solo: vuol dire che tutti questi infiniti vettori sono tra loro proporzionali – ad esempio, $\mathbf{w} = 85/90\mathbf{v}$ – e quindi formano un sottospazio di dimensione 1, l'autospazio V_2 . Ha importanza dire che proprio \mathbf{v} , o proprio \mathbf{w} , è un autovettore? Non molta: sia \mathbf{v} che \mathbf{w} possono essere una base di V_2 , quindi l'uno vale l'altro. Scegliere un autovettore tra quegli infiniti vettori proporzionali equivale in tutto e per tutto a scegliere una base per uno spazio vettoriale; scegliamo il più semplice, $(0, 0, 1)$, e non ci pensiamo più.

Proviamo ora a usare altri vettori.

$$\text{se } \text{Paolo} = (42, 42, 90), \quad \text{allora } F(\text{Paolo}) = (42, 42, 180)$$

$$\text{se } \text{Maria} = (40, 40, 170), \quad \text{allora } F(\text{Maria}) = (40, 40, 340)$$

F è quindi un operatore ‘crescita’ che fa aumentare la statura di Paolo e di Maria, prima ragazzini poi adolescenti, lasciando invariata la loro linea.

Esempio 3.2. Sia $G(x, y, z) = (2x, 2y, z)$. Prendiamo i vettori $\mathbf{v} = (42, 42, 0)$, $\mathbf{w} = (40, 40, 0)$, e proviamo:

$$G(42, 42, 0) = (84, 84, 0) = 2(42, 42, 0)$$

$$G(40, 40, 0) = (80, 80, 0) = 2(40, 40, 0)$$

Chiaramente \mathbf{v} e \mathbf{w} sono autovettori di autovalore 2. Possiamo però trovarne anche altri un po' diversi:

$$G(42, 0, 0) = (84, 0, 0) = 2(42, 0, 0)$$

$$G(0, 40, 0) = (0, 80, 0) = 2(0, 40, 0)$$

Dopo l'esempio precedente la cosa non dovrebbe stupire: \mathbf{v} e \mathbf{w} sono vettori di un sottospazio di dimensione 2, una cui base semplice è $\{(1, 0, 0), (0, 1, 0)\}$. Proviamo ora a usare i vettori Paolo = (42, 42, 180) e Maria = (40, 40, 170):

$$\begin{aligned} \text{se } \text{Paolo} &= (42, 42, 180), \quad \text{allora } G(\text{Paolo}) = (84, 84, 180) \\ \text{se } \text{Maria} &= (40, 40, 170), \quad \text{allora } G(\text{Maria}) = (80, 80, 170) \end{aligned}$$

G appare così come un operatore 'ingrasso' che rovina la linea di Paolo e di Maria: la statura rimane la stessa (sono già cresciuti), ma larghezza e profondità aumentano al punto che urge una dieta.

Conclusione seria: se trovo autovettori, autovalori e autospazi relativi a un operatore lineare $T : V \rightarrow V$, posso scomporre il suo dominio V in sottospazi che variano indipendentemente da altri. La cosa è talmente utile che ha un nome preciso: *scomposizione di uno spazio vettoriale nella somma diretta di sottospazi invarianti*.

Per *sottospazio invariante* si intende appunto un sottospazio W di V tale che, applicando T a un qualsiasi suo elemento, si ottiene ancora un elemento di W .

Si dice anche che è possibile la *restrizione* di T al sottospazio invariante W , ottenendo così un operatore $T|_W : W \rightarrow W$. Ad esempio, si può restringere l'operatore F (quello dell'esempio 3.1) ai soli vettori del tipo $(0, 0, k)$.

In pratica, questo vuol dire che si può capire come funziona un operatore esaminando separatamente la sua azione sui diversi autospazi, uno per volta.

3.2 Autovalori e polinomi caratteristici

Primo problema: come trovare gli autovalori relativi agli autovettori di un operatore lineare? Normalmente si sa fare, ma ricapitoliamo brevemente.

Perché un λ sia un autovalore, deve valere $T(\mathbf{v}) = \lambda\mathbf{v}$, ovvero $T(\mathbf{v}) - \lambda\mathbf{v} = \mathbf{0}_V$.

Per poter andare oltre, devo trasformare $\lambda\mathbf{v}$ in qualcosa di simile a un'applicazione lineare; basta usare l'applicazione identità a cui, se le basi del dominio e del codominio sono uguali,⁹ è associata la matrice identità. Infatti $\lambda\mathbf{v}$ è evidente-

mente uguale a $\lambda Id_V(\mathbf{v})$; ad esempio, se $\lambda = 2$ e $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\lambda\mathbf{v} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$ è uguale a

$$2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Partendo da $T(\mathbf{v}) - \lambda Id_V(\mathbf{v}) = \mathbf{0}_V$, 'metto in evidenza' \mathbf{v} e ottengo:

$$(T - \lambda Id_V)(\mathbf{v}) = \mathbf{0}_V$$

Eseguo ora un passaggio che *sembra* facile e scontato, ma in realtà andrebbe capito bene in tutte le sue implicazioni: sostituisco T con una matrice $n \times n$ ad esso associata rispetto a una base B , che indico con \mathbf{A} , e Id_V con la matrice identità \mathbf{I}_n , dove n è la dimensione di V :

$$(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x} = \mathbf{0}_V$$

⁹Normalmente si usano sempre basi uguali per il dominio e per il codominio quando si lavora con operatori lineari, cioè con applicazioni del tipo $T : V \rightarrow V$, in quanto dominio e codominio coincidono.

Da notare che al posto di \mathbf{v} c'è ora \mathbf{x} , che è il vettore delle coordinate di \mathbf{v} rispetto alla base B , perché una matrice – lo ripetiamo – opera sempre e solo su coordinate rispetto a una base.

A questo punto la strada è tutta in discesa:

- osservo che $(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x} = \mathbf{0}_V$ non è altro che un sistema lineare omogeneo;
- considero che la soluzione $\mathbf{x} = \mathbf{0}_V$ non interessa, in quanto sto cercando autovettori e questi, per definizione, non possono essere nulli;
- ricordo che, affinché un sistema lineare omogeneo abbia soluzioni diverse da quella nulla, il determinante della matrice dei coefficienti deve essere 0;
- la matrice dei coefficienti è qui $(\mathbf{A} - \lambda \mathbf{I}_n)$; contiene un'incognita, λ , quindi non posso calcolare direttamente il determinante, ma posso solo esprimerlo in forma algebrica;
- ottengo così il cosiddetto **polinomio caratteristico**, un polinomio in λ , e ne cerco le radici, cioè i valori di λ per cui quel polinomio (quindi il determinante) si annulla.

Gli autovalori che cerco sono appunto le radici del polinomio caratteristico.

Sono solo conti, quindi in genere si sanno fare. C'è solo da dare un consiglio: *non è necessario, e nemmeno opportuno, arrivare davvero al polinomio facendo tutti i conti; è spesso meglio fermarsi non appena si trovano fattori 'semplici' del polinomio.*

Esempio 3.3. Sia $\mathbf{A} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. Allora $(\mathbf{A} - \lambda \mathbf{I}_3) = \begin{bmatrix} -\lambda & -1 & 0 \\ 0 & 1 - \lambda & 0 \\ 1 & 1 & 1 - \lambda \end{bmatrix}$.

Scelgo la seconda riga per il calcolo del determinante e svolgo tutti i passaggi:

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}_3) &= (1 - \lambda) \begin{vmatrix} -\lambda & 0 \\ 1 & 1 - \lambda \end{vmatrix} = \\ &= (1 - \lambda)[- \lambda(1 - \lambda)] = \\ &= -\lambda(1 - \lambda)^2 = \\ &= -\lambda(1 - 2\lambda + \lambda^2) = \\ &= -\lambda + 2\lambda^2 - \lambda^3 \end{aligned}$$

Sono stato 'bravo', ho ottenuto un 'vero' polinomio, ma è di terzo grado... Per venirne a capo dovrei scomporre il polinomio in fattori, ma se mi fermo prima, se cioè mi fermo a $-\lambda(1 - \lambda)^2$, ho già la scomposizione in fattori e vedo subito che il polinomio si annulla per $\lambda = 0$ e per $\lambda = 1$.

Ci sarebbero a questo punto diversi aspetti importanti da sottolineare, che si capirebbero meglio proponendo le relative dimostrazioni. Dato che però le dimostrazioni si dimenticano presto... mi limito a un elenco:

- 1) Il polinomio caratteristico di una matrice quadrata di ordine n è sempre un polinomio di grado n .
- 2) Trovare le radici di un polinomio p di grado n vuol dire risolvere l'equazione $p = 0$. Come ben noto, un'equazione di grado n ha esattamente n soluzioni reali

o complesse, ma può ben avere meno di n soluzioni reali. Ad esempio, l'equazione $x^2 + 1 = 0$ ha due soluzioni complesse ma nessuna soluzione reale. Se quindi ci si limita alle soluzioni reali, può ben capitare di trovare meno di n soluzioni dell'equazione $p = 0$, quindi meno di n radici del polinomio p . *In linea di principio, quindi, ci sono tanti autovalori quanta è la dimensione del dominio dell'operatore, di V , ma se ci si limita a quelli reali se ne possono trovare meno.*

- 3) Nell'esempio abbiamo trovato 2 soluzioni, 0 e 1, per l'equazione di terzo grado $-\lambda + 2\lambda^2 - \lambda^3 = 0$, ma in realtà sono tutte. Questo si spiega perché abbiamo scomposto il polinomio in due fattori e trovato quindi le soluzioni delle equazioni $-\lambda = 0$ e $(1 - \lambda)^2 = 0$. Nel primo caso è evidente che c'è un'unica soluzione; quanto al secondo, se si prova a svolgere il quadrato si trova:

$$\lambda^2 - 2\lambda + 1 = 0$$

che si risolve con la formula $(2 \pm \sqrt{4 - 4})/2$, il cui discriminante è nullo; si dice in questi casi che l'equazione ha due soluzioni coincidenti, oppure anche – ed è questa la terminologia che si adotta quando si lavora sugli autovalori – che la soluzione 1 ha **molteplicità algebrica** 2; si scrive $m_a(1) = 2$. In generale, *una volta scomposto il polinomio caratteristico in fattori, le sue radici hanno molteplicità algebrica pari al loro esponente*; ad esempio, se ho il polinomio

$$\lambda(\lambda - 1)^2(\lambda + 2)^3(\lambda - 5)^4$$

le sue radici sono: 0 con $m_a(0) = 1$, 1 con $m_a(1) = 2$, -2 con $m_a(-2) = 3$, 5 con $m_a(5) = 4$. Questo era un caso di polinomio con radici tutte reali. In casi analoghi, il numero totale delle radici trovate è uguale al grado del polinomio caratteristico, quindi all'ordine della matrice, ed è anche uguale alla somma delle molteplicità algebriche. Se il polinomio fosse stato:

$$\lambda(\lambda^2 + 1)^2(\lambda + 2)^3(\lambda - 5)^4$$

il secondo fattore non avrebbe avuto radici reali e la somma delle molteplicità algebriche delle radici reali sarebbe stata minore dell'ordine della matrice (quindi minore della dimensione dello spazio V).

- 4) *La traccia di una matrice \mathbf{A} , cioè la somma degli elementi della sua diagonale principale, è uguale alla somma dei suoi autovalori, ciascuno considerato con la sua molteplicità algebrica.* Ad esempio, abbiamo visto che gli autovalori della matrice \mathbf{A} sono 0 con $m_a(0) = 1$ e 1 con $m_a(1) = 2$, quindi la loro somma, tenendo conto della molteplicità algebrica, è $0 + 1 + 1 = 2$, che è anche la traccia di \mathbf{A} .

- 5) *Il determinante di una matrice \mathbf{A} è uguale al prodotto dei suoi autovalori.* Nel nostro caso, il prodotto degli autovalori è $0 \cdot 1 \cdot 1 = 0$, come il determinante di \mathbf{A} .

È evidente quanto le ultime due proprietà possano risultare utili per verificare se si sono calcolati correttamente gli autovalori...

Osservazione. Vorrei tanto che nella mente di chi legge si affacciasse un dubbio: ma se per trovare gli autovalori di un'operatore devo lavorare sulla matrice ad esso associata *rispetto a una base*, chi mi dice che cambiando base non troverei autovalori

diversi?

Quanto abbiamo già visto sulle matrici simili consente di sciogliere il dubbio. Abbiamo visto che due matrici \mathbf{A} e \mathbf{B} sono simili se si può trovare una matrice invertibile \mathbf{N} tale che $\mathbf{B} = \mathbf{N}^{-1}\mathbf{A}\mathbf{N}$. Abbiamo anche visto che *matrici simili rappresentano uno stesso operatore lineare rispetto a basi diverse*.

Basta quindi dimostrare che *matrici simili hanno lo stesso polinomio caratteristico, quindi gli stessi autovalori (quindi la stessa traccia, quindi lo stesso determinante)*, e infatti:

$$\begin{aligned}\det(\mathbf{B} - \lambda \mathbf{I}_n) &= \det(\mathbf{N}^{-1}\mathbf{A}\mathbf{N} - \lambda \mathbf{I}_n) = \det(\mathbf{N}^{-1}\mathbf{A}\mathbf{N} - \mathbf{N}^{-1}(\lambda \mathbf{I}_n)\mathbf{N}) = \\ &= \det(\mathbf{N}^{-1}(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{N}) = \det(\mathbf{N}^{-1}) \det(\mathbf{A} - \lambda \mathbf{I}_n) \det(\mathbf{N}) = \\ &= \det(\mathbf{A} - \lambda \mathbf{I}_n)\end{aligned}$$

Lo sviluppo sulla prima riga si basa sul fatto che la matrice identità è simile a se stessa: $\mathbf{N}^{-1}\mathbf{I}_n\mathbf{N} = \mathbf{N}^{-1}\mathbf{N} = \mathbf{I}_n$.

Lo sviluppo sulla seconda si basa sul fatto che il determinante del prodotto di due o più matrici è uguale al prodotto dei relativi determinanti.

La conclusione si basa sul fatto che il determinante di una matrice e quello della sua inversa sono l'uno il reciproco dell'altro, quindi $\det(\mathbf{N}^{-1}) \det(\mathbf{N}) = 1$.

Va ricordato e lo ripeto: *matrici simili hanno lo stesso polinomio caratteristico, quindi gli stessi autovalori, quindi la stessa traccia, quindi lo stesso determinante*.

3.3 Autovettori e autospazi

Se si è capito cosa sono in generale autovalori, autovettori e autospazi, se si è imparato a trovare gli autovalori, individuare autovettori e autospazi non dovrebbe presentare alcun problema.

Ricordiamo: dato un operatore lineare $T : V \rightarrow V$, scelta una matrice associata \mathbf{A} rispetto a una base B , si trovano scalari λ tali che $T(\mathbf{v}) = \lambda \mathbf{v}$, nonché $\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$, dove \mathbf{x} è il vettore delle coordinate di \mathbf{v} rispetto a B . Può così succedere che V risulti scomponibile in autospazi 'indipendenti' (detti *sottospazi invarianti*), ciascuno dei quali relativo a un autovalore. Per autovettori relativi a λ si intendono gli infiniti vettori del corrispondente autospazio, rappresentati sinteticamente da una base di questo.

In sostanza, si procede così: autovalori \rightarrow autospazi \rightarrow autovettori, e basta applicare quello che già si sa.

Esempio 3.4. Sia $T : \mathbb{R}_2[t] \rightarrow \mathbb{R}_2[t]$ definita da:

$$T(1) = t^2, \quad T(t) = -1 + t + t^2, \quad T(t^2) = t^2$$

con base $B = \{1, t, t^2\}$.

La matrice associata a T rispetto a B si trova subito, in quanto abbiamo le immagini degli elementi della base:

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Per trovare gli autovalori esaminiamo il polinomio caratteristico della matrice

$$(\mathbf{A} - \lambda \mathbf{I}_3) = \begin{bmatrix} -\lambda & -1 & 0 \\ 0 & 1 - \lambda & 0 \\ 1 & 1 & 1 - \lambda \end{bmatrix}$$

e troviamo, come abbiamo già visto nell'esempio 3.3, gli autovalori $\lambda_1 = 0$ con $m_a(0) = 1$ e $\lambda_2 = 1$ con $m_a(1) = 2$.

A questo punto basta sostituire in $(\mathbf{A} - \lambda \mathbf{I}_3)$ i valori 0 e 1 a λ :

a) per $\lambda = 0$, l'autospazio V_0 non è altro che lo spazio delle soluzioni del sistema $(\mathbf{A} - 0\mathbf{I}_3)\mathbf{x} = \mathbf{A}\mathbf{x} = \mathbf{0}$:

$$\text{Sol}(\mathbf{A}, \mathbf{0}) = \left\{ k \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} : k \in \mathbb{R} \right\}$$

ricordiamo peraltro che le matrici operano solo su coordinate e che quindi, lavorando su una matrice, abbiamo ottenuto un vettore di coordinate che dobbiamo convertire in un polinomio; una base di V_0 è quindi il polinomio $\mathbf{v}_1 = (1 - t^2)$;

b) per $\lambda = 1$, l'autospazio V_1 non è altro che lo spazio delle soluzioni del sistema $(\mathbf{A} - 1\mathbf{I}_3)\mathbf{x} = (\mathbf{A} - \mathbf{I}_3)\mathbf{x} = \mathbf{0}$:

$$\text{Sol}(\mathbf{A} - \mathbf{I}_3, \mathbf{0}) = \left\{ h \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} : h, k \in \mathbb{R} \right\}$$

convertendo le coordinate in polinomi otteniamo $\mathbf{v}_{21} = (1 - t)$ e $\mathbf{v}_{22} = t^2$, ovvero $\{1 - t, t^2\}$ come una base di V_1 .

La dimensione dell'autospazio V_{λ_i} relativo a un autovalore λ_i viene detta **molteplicità geometrica** dell'autovalore e si indica con $m_g(\lambda_i)$.

Dato che normalmente si sa trovare lo spazio delle soluzioni di un sistema lineare omogeneo, non ci dilunghiamo. Vorrei piuttosto chiarire un possibile dubbio.

Si dimostrano facilmente due importanti teoremi:

1) *la molteplicità geometrica di un autovalore è sempre maggiore o uguale a 1 e minore o uguale alla sua molteplicità algebrica:*

$$1 \leq m_g(\lambda_i) \leq m_a(\lambda_i)$$

questo vuol dire che, in un caso simile a quello dell'esempio, avremmo potuto trovare un autospazio V_1 di dimensione 1 (che è la dimensione minima) nonostante la molteplicità algebrica dell'autovalore $\lambda_2 = 1$ sia 2;

2) *autovettori relativi ad autovalori distinti sono linearmente indipendenti*; in genere si dice: se si hanno k autovalori distinti $\lambda_1, \lambda_2, \dots, \lambda_k$, i relativi autovettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ sono linearmente indipendenti.

Per quanto ovvio, non sempre si aggiunge che è linearmente indipendente l'insieme di *tutti* gli autovettori anche se si considerano autovalori *non* distinti, cioè con molteplicità algebrica maggiore di 1. Se λ_1 ha molteplicità algebrica e geometrica

maggiore di 1, allora saranno più di uno anche i relativi autovettori; questi saranno tra loro linearmente indipendenti (costituiscono una base di un autospazio) e risulteranno linearmente indipendenti anche da autovettori relativi ad altri autovalori. Ad esempio, se il dominio V ha dimensione 5 e se si trovano 5 autovalori non tutti distinti come:

$$\lambda_1, \lambda_1, \lambda_1, \lambda_2, \lambda_2 \quad \text{con } m_a(\lambda_1) = 3, m_a(\lambda_2) = 2$$

se inoltre sia λ_1 che λ_2 hanno molteplicità geometrica 2, si avranno 4 autovettori:

$$v_{11}, v_{12}, v_{21}, v_{22}$$

e questi risulteranno linearmente indipendenti anche se relativi ad autovalori non tutti distinti.

In pratica, questo vuol dire che *se la somma delle dimensioni degli autospazi, quindi la somma delle molteplicità geometriche degli autovalori, è uguale alla dimensione del dominio V dell'operatore, l'unione delle basi degli autospazi costituisce una base di V* , e V può essere quindi scomposto nella somma diretta dei suoi autospazi.

3.4 Diagonalizzazione di un operatore lineare

Qui se ne vedono di tutti i colori. Agli esami.

L'errore più frequente consiste nel rispondere a una domanda sulla *definizione* di operatore diagonalizzabile enunciando il *criterio* di diagonalizzabilità. Come se alla domanda «cosa vuol dire matrimonio?» si rispondesse «non essere sposati». Distinguiamo quindi bene:

Operatore diagonalizzabile. *Un operatore lineare $T : V \rightarrow V$ tale che esista una base N di V costituita da autovettori per T .*

Viene poi fuori che la matrice associata a T rispetto a una base di autovettori è una matrice diagonale (vedremo subito perché).

Condizioni di diagonalizzabilità. *Un operatore lineare $T : V \rightarrow V$ è diagonalizzabile se e solo se:*

- a) la somma delle molteplicità algebriche dei suoi autovalori è uguale alla dimensione di V ;*
- b) per ogni autovalore, molteplicità algebrica e molteplicità geometrica coincidono.*

Esamineremo prima le condizioni, poi torneremo sulla definizione.

Da quanto già detto nei messaggi precedenti, dovrebbero risultare chiari i motivi per cui le due condizioni sono entrambe necessarie (e sufficienti).

Cominciamo dalla seconda:

- a) gli autovettori, in quanto elementi di autospazi, sono infiniti (trovatone uno, basta moltiplicarlo per un qualsiasi scalare per averne un altro); quando si parla di autovettori in numero finito, si intende riferirsi a quelli che costituiscono possibili basi degli autospazi;

- b) l'unione di basi degli autospazi è sempre un insieme linearmente indipendente di vettori e, pertanto, *può* costituire una base di V a condizione che la somma delle dimensioni degli autospazi, quindi delle molteplicità geometriche degli autovalori, sia uguale alla dimensione di V ;
- c) affinché questo succeda, la somma delle molteplicità geometriche deve essere uguale a quella delle molteplicità algebriche, in quanto:
- per ciascun autovalore la molteplicità geometrica può essere inferiore o uguale, ma non superiore, alla molteplicità algebrica e, se è inferiore, si ha inesorabilmente $\sum_i m_g(\lambda_i) < \dim V$;
 - la somma delle molteplicità algebriche, che altro non è che il grado del polinomio caratteristico, è uguale alla dimensione di V ... se si tengono in conto anche autovalori complessi.

Quest'ultima considerazione ci riporta alla prima condizione.

Come abbiamo già notato (sez. 3.2), se ci limitiamo ad autovalori reali può succedere che ne troviamo meno del grado del polinomio caratteristico. In questo caso la somma delle molteplicità algebriche risulta inferiore alla dimensione di V e quindi la somma delle molteplicità geometriche è comunque insufficiente, anche se si rispetta la seconda condizione.

In altri termini, la prima condizione serve solo se ci si limita ad autovalori reali, è invece automaticamente soddisfatta se si considerano anche eventuali autovalori complessi.

Vediamo ora *perché la matrice associata a un operatore rispetto a una base di autovettori è diagonale*.

Partiamo dal solito operatore $T : V \rightarrow V$ e fissiamo per V una base N costituita da autovettori.

La matrice associata a T rispetto alla base N deve avere per colonne le coordinate rispetto a N delle immagini degli elementi di N .

Indichiamo con \mathbf{L} la matrice associata, con \mathbf{L}_j la sua j -esima colonna, con \mathbf{v}_j il j -esimo autovettore, con λ_j il corrispondente autovalore. Si ha: $T(\mathbf{v}_j) = \lambda_j \mathbf{v}_j$.

Per trovare \mathbf{L} si devono individuare le coordinate dei \mathbf{v}_j rispetto a N , che sono però semplicemente $(0, \dots, 1, \dots, 0)$, con tutti zeri tranne un 1 al j -esimo posto. Ad esempio, se $N = \{(1, 0, -1), (1, -1, 0), (0, 0, 1)\}$, le coordinate del secondo autovettore rispetto a N sono evidentemente $(0, 1, 0)$ e le coordinate della sua immagine sono $(0, \lambda_2, 0)$.

Ne segue che la j -esima colonna di \mathbf{L} sarà:

$$\mathbf{L}_j = \text{Coord}_N(\lambda_j \mathbf{v}_j) = \lambda_j \text{Coord}_N(\mathbf{v}_j) = \begin{bmatrix} 0 \\ \dots \\ \lambda_j \\ \dots \\ 0 \end{bmatrix}$$

con tutti zeri tranne λ_j al j -esimo posto.

Esempio 3.5. Torniamo all'esempio 3.4, con un operatore $T : \mathbb{R}_2[t] \rightarrow \mathbb{R}_2[t]$ definito da:

$$T(1) = t^2, \quad T(t) = -1 + t + t^2, \quad T(t^2) = t^2$$

La matrice associata a T rispetto alla base $B = \{1, t, t^2\}$ è:

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Il suo polinomio caratteristico, cioè il determinante della matrice $(\mathbf{A} - \lambda \mathbf{I}_3)$, è $-\lambda(1 - \lambda)^2$. Le sue radici sono gli autovalori $\lambda_1 = 0$ con $m_a(0) = 1$ e $\lambda_2 = 1$ con $m_a(1) = 2$. Prendo nota: la somma delle molteplicità algebriche è uguale a 3, che è anche la dimensione di $\mathbb{R}_2[t]$.

Gli autospazi sono V_0 , una cui base è $\{1 - t^2\}$, e V_1 , una cui base è $\{1 - t, t^2\}$.

La somma delle molteplicità geometriche, cioè delle dimensioni degli autospazi, è uguale alla somma delle molteplicità algebriche, quindi alla dimensione di $\mathbb{R}_2[t]$. Posso quindi dire che

$$N = \{1 - t^2, 1 - t, t^2\}$$

è una base di autovettori per T dello spazio vettoriale.

A questo punto diagonalizzo T , cioè trovo una matrice diagonale ad esso associata rispetto alla base N ; per quanto appena visto, questa sarà una matrice diagonale con gli autovalori sulla diagonale principale:

$$\mathbf{L} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Si verifica facilmente che \mathbf{A} e \mathbf{L} sono simili, cioè che, indicando con \mathbf{N} la matrice avente per colonne gli autovettori per T (elementi della base N), si ha:

$$\mathbf{A} = \mathbf{N} \mathbf{L} \mathbf{N}^{-1} \quad \text{e} \quad \mathbf{L} = \mathbf{N}^{-1} \mathbf{A} \mathbf{N}$$

Possiamo anche verificare che, in quanto simili, le due matrici rappresentano l'operatore T rispetto a due basi diverse, \mathbf{A} rispetto a B , \mathbf{L} rispetto a N ; ad esempio, prendendo $p = 1 + t + t^2$:

a) $T(p) = -1 + t + 3t^2$;

b) con $\mathbf{A} \text{Coord}_B(p) = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}$ ottengo le coordinate rispetto a B di $-1 + t + 3t^2$;

c) le coordinate rispetto a N di p sono $(2, -1, 3)$ e con:

$$\mathbf{L} \text{Coord}_N(p) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}$$

ottengo le coordinate rispetto a N di $-1 + t + 3t^2$, infatti:

$$0(1 - t^2) - 1(1 - t) + 3(t^2) = -1 + t + 3t^2$$

Osservazione. Trovati gli autospazi rispetto a T , ho scomposto $\mathbb{R}_2[t]$ nella somma diretta di due sottospazi che seguono regole diverse, cioè di due sottospazi invarianti (cfr. sez. 3.1). In particolare:

- a) il sottospazio costituito dai multipli del polinomio $1 - t^2$, essendo 0 il relativo autovalore, non è altro che il nucleo dell'applicazione; infatti:

$$T(k(1 - t^2)) = kT(1 - t^2) = k[T(1) - T(t^2)] = k[t^2 - t^2] = 0$$

- b) gli elementi del sottospazio costituito dalle combinazioni lineari dei polinomi $1 - t$ e t^2 rimangono immutati in quanto il relativo autovalore vale 1; infatti, indicando la generica combinazione con $h(1 - t) - kt^2$:

$$T(h - ht - kt^2) = ht^2 - h(-1 + t + t^2) - kt^2 = ht^2 + h - ht - ht^2 - kt^2 = h - ht - kt^2$$